VOL. XVI JUNE, 1932 No. 3

The Journal of
APPLIED PSYCHOLOGY

JAMES P. PORTER, EDITOR
Ohio University

WITH A BOARD OF CO-OPERATING EDITORS

TABLE OF CONTENTS

JOHN RUSKIN DYER. Sources and Permanence of Vocational Interests of College Men 233
DONALD A. LAIRD. How the Consumer Estimates Quality by Subconscious Sensory Impressions 241
EDWARD L. THORNDIKE. The Significance of Responses in the Free Association Test 247
DAVID WECHSLER. Analytic Use of the Army Alpha Examination 254
M. N. THISTED and H. H. REMMERS. The Effect of Temporal Set on Learning 257
G. R. WENDT. Two Industrial Applications of a Psychophysical Method 269
ROY M. DORCUS. Habitual Word Associations to Colors as a Possible Factor in Advertising 277
HERBERT L. BOWMAN. A Technique for Determining the Relationship between Reported Preference and Performance 288
MAX L. HUTT. The Kohs Block-Design Test: A Revision for Clinical Practice 298
KEY LEE BARKLEY. A Consideration of the Differences in Readiness of Recall of the Same Advertisements by Men and Women 308
JOHN W. CHARLES. Competition for Marks 315
J. C. LAUDERBACH and ENID HAUSE. The Reliability and Validity of Derived Scores Yielded by the McCall Multi-Mental Scale 322
PAUL HORST. A Routine Procedure for Obtaining Comparable Scores 324
DOROTHY DURLING. Note on the Comparative Reliability of the Stanford-Binet Below the Age of Six 331
Notes and News 334
Book Reviews 337
New Books and Pamphlets 341

PUBLISHED BI-MONTHLY BY THE EDITOR
Ohio University, Athens, Ohio

Printed by
THE SCIENCE PRESS PRINTING COMPANY
LIME AND GREEN STS., LANCASTER, PA.

Entered as second-class matter at the post office at Athens, Ohio, under the act of March 3, 1879. Additional entry as second-class matter at the post office, Lancaster, Pa. Acceptance for mailing at special rate of postage provided for in section 1103, act of October 3, 1917. Authorized December 21, 1922.

COPYRIGHT 1932, BY JAMES P. PORTER

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. F. Book, Indiana University, W. V. Bingham, Personnel Research Federation, New York; June E. Downey, University of Wyoming; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; J. E. W. Wallin, Atlantic University; Herbert S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College; Henry M. McComas, Princeton University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, Boston, Mass.; Edward K. Strong, Jr., Leland Stanford Jr. University; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; Joseph Jastrow, New School of Social Research, New York; G. M. Whipple, Danvers, Mass.; Helen Thompson Wooley, Teachers College, Columbia University.

The scope of The Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeals. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or playing musical instruments, sports, etc.

Communications, manuscripts, reviews, summaries, notices, books and off-prints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by this Journal; in every case the year, volume and pages should be specified.

The Journal of Applied Psychology is published bi-monthly. The subscription price is $5.00 a year. Single numbers $1.00 to $2.00 depending upon the number of pages. Remittances and business communications should be addressed to the Editor, James P. Porter, Ohio University, Athens, Ohio.

SOURCES AND PERMANENCE OF VOCATIONAL INTERESTS OF COLLEGE MEN—101 CASES OVER FIVE YEAR PERIOD

JOHN RUSKIN DYER

University of Idaho, Southern Branch

In the spring of 1924 the writer asked one hundred one students at the University of Kansas to come to his office for conferences on their vocational situations. They were boys with whom he was personally acquainted and the conferences were quite informal. They were nearly all juniors, seniors, or graduate students.[1]

This initial statement was made to each of them—"I am interested in vocational guidance for college students and want you to help me get a good picture of the situation with a group of our own college students. I need your help. Will you kindly tell me as fully as you can just how you have reached your own choice of a vocation? Begin as far back as you can remember and tell the story in your own way without interruption from me. I want to make notes while you talk. You may be sure that your identity will not be revealed in the publication of the study and that you will be protected from every possible embarrassment in the use of your story."

Thereupon each student told his story in his own words. It usually required from five to twenty minutes for the personal statement although a few took much longer. When this statement was completed the writer glanced over the following items to see if they were included in the informal statement. If not, questions were asked to bring them into the report.

[1] The recent survey of the literature on interest in Fryer, D., "The Measurement of Interest," 1931, makes unnecessary an historical introduction to this paper.

1. The family situation with reference to vocational opportunities for the boy.
2. The kind of community or communities in which the boy had lived—whether rural, or urban, etc.
3. The boyhood interests and hobbies.
4. All occupations at which the boy had worked.
5. The favorite studies in High School and College.
6. High school, college, and community "Activities."
7. The most interesting and enjoyable experiences in his life.
8. The boy's own definition of success at age fifty.
9. The sources of vocational counsel received.
10. The lifework choices in order of preference.
11. The vocational plans for the following ten years if the boy were to be given $100,000 for the conference.

The most interesting part of the study is the one hundred one life histories. It is, of course, impossible to print them with this article. They are included, however, with the report in the library of the Graduate School of the University of Kansas. There are stories of heroism, inspiration, careful planning and often of drifting in these life histories.

Number one, after working his way through high school, went to college to continue the job. To make it all the more difficult he married during his second year at college and became a father the third year. After trying clerking and inspecting, both white collar occupations, he turned to janitor work as more remunerative and spent the last three of his five college years that way. At *one time* he held fifteen different jobs totalling $165 a month. He arose every morning at four forty-five and retired with the furnaces at night. But he completed the educational job, secured a good position, has had one unusual promotion and is on his way up.

Number eight ran into everything imaginable that could kill his enthusiasm for college, but simply refused to quit. An extra year of high school because of bad advice, the World War, an injury in the service, at least seven different attempts at college, lack of funds, and still he stayed with it. An interesting slant in his case is that his love of medicine was

originally vicarious—his father's boyhood ambition had been to be a doctor but he had given it up. The son inherited that dream and eventually made it come true.

Thirty-two of the one hundred one made their college choice of a vocation before reaching high school; twenty-eight made the decision in high school; thirty-seven made it in college. Two reached their decisions on the job, one at age seventeen and the other at age twenty-three. In two cases it was not at all clear when the decisions were made. In some doubtful cases the time of decision was recorded as "probably" within a given period.

The origin of the choice was not always clear. In thirty-one cases it seemed to be what the writer calls the "family situation or tradition." These cases seemed to follow a father or other close relative's profession or business often with the prize of a partnership as a compelling motive. Again certain doubtful cases are recorded "probably."

In fourteen cases a boyhood occupation was made the permanent choice. A common example was the trend from a paper route to journalism. In ten cases a hobby became the beginning for a real lifework. An example would be the transition from playing with electrical toys to the vocation of electrical engineering. Twelve boys seemed to base their decisions on the influence of a teacher, aside from formal counsel, or in a course of study. This happened in both high school and college. Nine boys were guided by formal counsel from parents, other relatives, teachers and schoolmates.

The Permanence of Vocational Interests Five Years Later

When the study thus outlined was completed it was laid aside for several years when it seemed wise to revive it. Accordingly information was secured telling what occupations the boys actually entered and how long they remained in each one. All of the boys were still living and the last of their reports came in recently showing the vocational selections for approximately five years following the original study.

TABLE 1
Source of first choices
11 cases

COLLEGE MAJOR		TIME OF CHOICE		APPARENT ORIGIN OF CHOICE	
Business	31	Before High School	32	Family situation or occupation	31
Engineering	16	During High School	28	Boyhood occupations	14
Law	12	During College	37	Classroom Inspiration	12
Journalism	12	On the Job	2	Boyhood Hobbies	10
Medicine	7	Not Clear	2	Counsel of Friends	9
Education	3			Origin Not Clear	15
Scattering	20			Scattering Origins	10
Total	101	Total	101	Total	101

This supplement may prove to be more valuable than the original study itself since it gives an authentic and objective report on the permanence of vocational interest.

One of the students revealed the weakness in a memory report of earlier interests by what lay beneath a very black smudge on his final report. When, through the curiosity of the writer, it was finally erased enough to reveal the sentence it was meant to conceal, these words appeared, "Dear Dean, whatinell did I say I wanted to do?" And this uncertainty despite the fact that he had never wavered in his original choice. Another phrased the same idea somewhat more conventionally when he wrote, "I should surely be interested in knowing what I thought I wanted to be when I talked to you. *I've forgotten that* but I do remember the interview. Can you satisfy my curiosity?" He, too, was following the program outlined five years previously. These two comments from the boys are reminiscent of Thorndike's own comment as to the chief weakness in his study.

The offer of the $100,000, and the request for a revision of plans if that much money were available brought a smile and

in a number of cases a real revision. Four would have shifted to Choice No. 2, two to Choice No. 3, and fourteen to a choice not even indicated earlier. Writing, travel, and law were mentioned often as goals not possible without additional financial help.

Eighty-two of the one hundred and one boys really first entered upon their indicated first choices. Three others entered into related fields and one eventually came back to his first choice after trying Choice No. 4. So that eighty-six of the one hundred one really gave their first choices a tryout. However, nine others first began with Choice No. 2 and six began with a choice not even indicated earlier. Seven of them changed to Choice No. 2 or to an unindicated choice while only one needed a third choice within the five years.

At the end of the period of approximately five years, seventy-nine were engaged at Choice No. 1, ten at Choice No. 2, three at Choice No. 3, and nine at vocations not listed as interesting them at all five years earlier.

Both the numbers actually entering upon their first choices and the numbers actually engaged therein five years after indicating their selections, which usually meant from three to five years after graduation, indicate a high degree of permanency in vocational interest as expressed in the college years.

Of the thirty-two who made their decisions before reaching high school twenty-nine or ninety percent were still true to their choices at the end of the period and the two who wavered went into business with members of their own families, and one of them turned to his $100,000 choice. One is reminded of the old adage, "Train a child up in the way he should go and when he is old he will not depart therefrom." Those who made their decisions in high school and college felt freer to change. Only seventy-five percent of the high school decisions and sixty-eight percent of those made in college were still effective at the close of the period studied.

Of the thirty-one who followed the easy route of the family tradition twenty-six or eighty-four percent remained faithful to the end. Of the twelve who found their inspiration in teacher or classroom nine or seventy-five percent stuck it out. Of the nine who listened to counsel from relatives, teachers, and students six or sixty-seven percent found it to their liking.

Of real significance are the two groups who got their vocational beginnings in an occupation or a hobby. Of the fourteen who dated their start from an occupation two changed and one of them really did not change the nature of his work. The other was influenced by the family situation and went to his $100,000 choice. Of the ten who found vocational guidance in a hobby nine held true while the one who left was simply the recipient of an offer too flattering to refuse and still

TABLE 2
The permanence of first and other choices
101 cases

WHAT CHANGE WITH GIFT OF $100,000		VOCATION ACTUALLY FIRST FOLLOWED ON LEAVING COLLEGE		VOCATION FOLLOWED AT END OF FIVE YEARS	
No change	81	First college choice	82	First college choice	79
To second choice	4	Second choice	9	Second choice	10
To third choice	2	Third choice	3	Third choice	3
To choice not previously mentioned	14	Fourth choice	1	Choice not previously mentioned	9
		Choice not previously mentioned	6		
Total	101	Total	101	Total	101

has the earlier interest hidden away in a corner of his heart. Thus occupations and hobbies represent almost one hundred percent holding power.

The study revealed no significant difference in the vocational holding power of different vocations. Medicine, law, engineering, business, journalism, and education, which were the larger groups, all suffered small losses but the percentages were not significantly different.

It would add an interesting chapter to the study to have the students concerned take the Stanford interest test for the choices indicated. These tests, of course, were still somewhat below the horizon in 1924.

TABLE 3

Permanence with reference to time of choice
Number in each group which held for five years

CHOICES MADE BEFORE HIGH SCHOOL	CHOICES MADE DURING HIGH SCHOOL	CHOICES MADE DURING COLLEGE
29* of 32 90%	21 of 28 75%	25 of 37 68%

* Two turned to the family occupation. The third to his $100,000 choice.

TABLE 4

Permanence with reference to source of choice
Number in each group which held for five years

FAMILY SITUATION	BOYHOOD OCCUPATION	CLASSROOM INSPIRATION	BOYHOOD HOBBIES	COUNSEL OF FRIENDS	ORIGIN NOT CLEAR	TOTAL GROUP
26 of 31 84%	12* of 14 86%	9 of 12 75%	9** of 10 90%	6 of 9 67%	9 of 15 60%	79 of 101 78%

* One who changed from first choice really did not change the nature of his work. The other turned to his $100,000 choice and the family occupation.

** The other was tempted by a flattering offer in another field and may return to his first choice.

General Conclusions

1. Vocational choices are made from the earliest years clear to the end of college. Vocational guidance should be an integral part of the entire educational system from kindergarten to graduate and professional schools as well as of the home life.

2. Hobbies and boyhood occupations are vital instruments of vocational selection.

3. It is likely that decisions made in line with the family tradition and the counsel of relatives, teachers, and friends should be subjected to unusually careful scrutiny.

4. With or without competent vocational guidance young folks make life work choices seriously enough to endow them with a high degree of permanence.

5. The college graduate in our complex modern life has a rich choice of vocations and can change from one to another with comparative ease.

6. The use of a permanent record card through all the school years could provide the information necessary for an extensive study of vocational interest.

7. If the Stanford Interest Test means what we hope for it in the high school and college years, it should be modified to serve the years below the high school where so many vocational decisions are made.

HOW THE CONSUMER ESTIMATES QUALITY BY SUBCONSCIOUS SENSORY IMPRESSIONS

With Special Reference to the Role of Smell

DONALD A. LAIRD*
Colgate University Psychological Laboratory

The consumer of manufactured articles rarely has a technical basis for judging the quality of the article which is to be purchased. Tensile strength, viscosity, wearing qualities and other essential features of quality which the specialist or the purchasing agent's laboratory can measure accurately cannot be determined by the usual retail purchaser. And yet the retail customer attempts to pick out a good quality, or the best quality, for the price.

It happens from time to time that an article which has excellent quality from the engineering point of view still does not impress the retail customer as possessing good quality, or that a poorly made article of poor materials impresses the customer as having excellent quality.

HOW IS QUALITY JUDGED?

What are some of the factors which influence the customer's judgment of quality? During a boom period recent history shows that the *desirability* of an article was judged in part by its color and design, and to some extent color and design were a part of the complex estimate of quality. Basic quality features, however, loom more important in present market conditions, and since the average purchaser has little technical skill in actually determining any real quality of manufactured

* With the assistance of Theodore Batsford, Garet P. Cowenhoven, Lloyd W. Cross, Morford Downs, Henry H. Graham, Donald Hodges, Alfred J. Jeffries, Richard J. Kehoe, W. Bartrum Jackson, E. B. MacSwan, Gordon Seymour, James Stephan, Emery W. Taylor.

articles it is important to learn something about the factors which must be active in influencing these judgments of quality.

It is apparent that these influences reach the customer through some sense. The senses most actively and consciously used by the customer are sight and touch; the fingers are run through a cloth, it is held up to the light and the customer looks through it, the dyeing is inspected for uniformity, it may be wadded lightly and held to the cheek to determine softness. The sense of taste or the muscular senses are more rarely used consiously.

TESTING THE ROLE OF SCENTS

The present experimental work is based upon the observation that the "general atmosphere" surrounding a product is often the determining element in the judgment of its quality. Packaging, store interior, prestige advertising are illustrations of this influence. The experimental work was done on the influence of faint scents upon the judgment of quality. Scents were chosen because competitive products vary markedly in this (due to different paints, finishing oils, etc.), because it was possible for us to vary the products in regard to their scent without altering their intrinsic engineering features and because observations of consumers suggested that the slight scent of a product was important in their selection of an article from a choice of several closely similar articles.

Women's silk hose were used for the experiment. They were bought in regular retail channels, were made by the same manufacturer, and were the same style, color, and design. So far as it is possible for textile machinery to make them, they were identical hose. Each pair of hose was packed in a separate box, and the boxes were identical.

A test set consisted of four pairs of these identical hose, packed in identical boxes. One pair was left with the natural slightly rancid scent which comes from the mixture of castor

oil and sulphates which is used in lubricating the fibers of the yarns to facilitate weaving and to yield softness. A second pair of hose was given a very faint scent by a compound of synthetic aromatic chemicals; the narcissus note predominated in this scent. A third pair of hose was given a complex fruity type of scent, and the fourth pair was given a sachet type of scent.

All these scents were very faint, and it was attempted to have them no more intense than the natural rancid smell of the pair that was not re-odorized; the aim was to change the quality or type of the scent rather than to change the strength of the scent. The reodorization was brought about by pinning a small card of perfumer's blotting paper to which a drop or aromatic compounds had been touched to each pair of hose; the aromatics could have been added to the finishing oils in the manufacture of the yarns, but since the synthetic aromatics had been compounded to blend with rather than to be stronger than the natural rancid smell, the use of the scented blotter produced the same aromatic results.

The scents on the reodorized hose were so faint that only six out of the 250 housewives tested noticed the scents. But the aromatic particles were present in the air entering the olfactory apparatus of the housewives and the data obtained show that through some subconscious channels their judgments of quality were influenced by this subconscious impression.

HOW THE DATA WERE OBTAINED

Men from the laboratory made a house-to-house survey of 250 housewives in Utica, New York, using a set of the four hose as described. Middle class and upper strata homes were visited. The men were instructed to use the following presentation:

"Good morning? I want you to help us find the best quality of silk stockings. I do not have a single thing to sell."

"Women who have had experience in selecting silk hosiery, who know quality, are helping us size-up quality in hosiery."

"Here are four pairs of hose that are very much alike. They are the same color, and all are made in exactly the same pattern and style. Which pair do you judge to be the best quality? Feel them in your fingers, look through them, stretch them, look at the seams. Do anything you would ordinarily do to pick out the best for your own use."

Do *not* tell or suggest to the lady that she smell the hose. Scatter the four opened boxes three or four feet apart, so that the scent on one pair does not spread to the others. The boxes can be opened and the covers placed beside them.

The blotter that is pinned to each pair should be kept with that pair, although the woman may want to take it off in order to run her fingers through the leg of the stocking. Watch closely to see that the hose and pinned cards are kept in the right box. The number, or letters, on the card and on the box should agree exactly.

Observe closely to note whether or not they smell the hose, and record this as well as other observations on the data blank *after* you are out of the house.

Any woman who will cooperate in the test can be used.

Be careful not to get tobacco smoke over the hose, or other smells mixed with the ones already placed in them.

In case the lady asks which *is* the best hose after her test is done, tell frankly that you do not know, and that the test is actually to find out which women judge to be the best. "Factory engineers' judgment is not to be counted, you see, madam, when the important thing is the experienced judgment of women themselves."

There was little difficulty in obtaining the cooperation of housewives until, fortunately, on the day when judgments from 250 women had been obtained. The experiment was discontinued on that day, January 13th, not due to any superstition connected with the date, but due to the publication of the following truthful news item in a local paper:

A scientific experiment in South Utica this forenoon alarmed a housewife and led her to telephone for the police.

The scare was caused by a group of Colgate University students posing as hosiery salesmen and bent on finding out whether or not a perfumed article of merchandise could be sold more readily than one that lacked the pleasant odor.

A student rang the bell of a home on Amy Avenue. No,

the woman who responded did not care to inspect the hosiery. Too busy.

"But would you just look at these stockings and tell me what you think of the quality," he persisted. Reluctantly she looked at the materials in two or three boxes and declared them all about the same.

The student was insisting that she should determine the quality, when the doorbell rang, indicating someone was at the rear door.

"Say, what's your racket?" asked the impatient housewife. "You get out of here, or I'll call the police." The young man left. "I'll answer one door at a time," said the woman and closed the door.

Going to the rear of the house she could find no one at the door. She became suspicious and called Lieut. James V. Felitto of the detective bureau. Detective Felitto soon had Detectives Thomas Ferrar and Chris McDerMott on the way to the house. They heard the story and started after the suspects.

In the Roosevelt Drive-Sunset Avenue section they came upon some Colgate students testing theories of salesmanship in the study of psychology.

Police learned the efforts were made at the suggestion of Prof. Donald Laird, professor of psychology at Colgate. They reported that the experimenters carried a box of three pairs of stockings which were perfumed and another with stockings that were without the artificial odor and their errand was to learn whether the housewives would consider the perfumed article more attractive than those without perfume.

The work was discontinued after the publication of this item, since it revealed to Utica housewives the purpose of the experiment and further data would have shown which scent they preferred rather than the rôle of subconscious sensory impressions upon the judgment of quality.

HOW HOUSEWIVES SAY THEY JUDGED THE BEST QUALITY

In the order of the occurrence, housewives thought they were determining the quality of the four identical pairs of hosiery by: texture, finer weave, feel, better wearing qualities, no sheen, heavier weight, and firmer weave. Although all pairs were the same, the housewives found no difficulty in telling

why one pair was the best, although some of them did have difficulty in finding out *which* pair was the best.

HOW THE HOUSEWIVES WERE INFLUENCED BY SCENTS IN JUDGING QUALITY

Each pair of hose should have received the same number of judgments as being the best quality, if the scents played no part in subconsciously influencing the judgment of quality. The following data which show that the scent played an important part in judging quality, have been arranged to show the reliability of the findings as well as the trend of the results:

SCENT	JUDGED THE BEST HOSIERY		
	1st 72 housewives	Next 83 housewives	All 250 housewives
"Natural"	8.3%	8.4%	8%
Sachet	16.6	18.0	18
Fruity	24.9	22.8	24
Narcissus	49.8	50.6	50

The statistical reliability of the data is shown by the close agreement of the first two columns with each other and with the records for the total group.

Had the data been gathered by one man we might justly suspect some forcing of the judgments in favor of the narcissus type of scent, but this is ruled out by the fact that more than a dozen men took about equal part in the data gathering, as well as by the fact that the men were prejudiced in advance by their personal favor for the sachet type of scent.

CONCLUSION

Two things are plainly indicated by this experiment: 1. Subconsciously perceived sensory impressions received through the olfactory apparatus are potent in determining the housewife's judgment of quality in silk hosiery. 2. Scents of one type are more influential in determining this judgment of quality than are scents of another type.

THE SIGNIFICANCE OF RESPONSES IN THE FREE ASSOCIATION TEST[1]

EDWARD L. THORNDIKE
Teachers College, Columbia University

In what follows I shall assume that the subject thinks of his task in the test as to the task of saying or writing the first word that the stimulus word calls up, or naming by a word the first thing or quality or event that it calls up.

For studies of individual eccentricity, peculiarities, interests, aversions, 'complexes' and the like, the free-association experiment with response by a word to a word is a useful instrument. But the reasons for its usefulness are not those which have been commonly assumed. The common assumption has been that the stimulus word makes a person think of the thing or quality or fact that it names, and that this thing or quality or fact calls up something else, the word for which the subject then writes. This is apparently very rarely the case, occurring in perhaps five per cent of Kent-Rosanoff responses. About ninety-five per cent of the responses are due to connections formed with the stimulus words in hearing and reading, and in speaking and writing. They reveal chiefly what the individual has heard, read, said and written, weighted by his interest therein, and what connections have operated to give meaning to words heard and seen.

The facts may be illustrated by the words *yours, good, cold, ja,* and *bed.* One of these words may (1) evoke some word which has been a frequent and satisfying accompaniment or sequent belonging to it in speech or writing, as in *yours truly, good morning, cold weather,* or *bed-clothes.* It may (2) evoke

[1] The facts reported here were obtained as a by-product of an investigation of learning, made with the aid of a grant from the Carnegie Corporation.

some word which has been a frequent and satisfying accompaniment or sequent, belonging to it in hearing or reading and helping to give it meaning, as in *yours* (*not*) *mine, good* (*not*) *bad, good nice, cold Alaska, cold frigid, cold* (*not*) *hot, ja yes, bed sleep.* It may (2a) evoke some non-verbal representation, which has been a past accompaniment of the word, of some thing or quality or event which helps to give meaning to the word. The subject may then write a word which names this non-verbal representation or is suggested by it. So we may have cold arousing wintry images and so the words, *snow, ice.* It may (2b) evoke some tendency which has functions like those just described but is non-representative, which tendency leads the subject to write its name or some word suggested by it. So cold may produce *b–r–r–r–r.* (3) The stimulus word may set in operation other connections than these of use and comprehension. So with *cold mold, cold pain,* and *cold yesterday.* I find the proportions of cases to be about 20 of (1), 75 of (2) and 5 or less of (3).[2] These estimates are derived from a detailed study of records obtained by myself from 162 or 192 educated adults, and of the records from 1000 children furnished by Woodrow and Lowell ['17], and from a less careful study of the records of Kent and Rosanoff ['27], and O'Connor ['28].[3]

The community of responses is in large measure due to the word-meaning connections. A person who oversteps this often may well be eccentric and unbalanced and with a sick brain. But this is not so much because he shows perversity in the concrete ideas which come when he hears the words, as because he shows derangement of the basal layer of connections by which words acquire meanings or direct speech.

[2] The argument is presented more fully and with evidence in Chapter 14 of The Fundamentals of Learning [Thorndike, '32]. Its starting-point was an investigation of the frequency of opposites among the responses. The connections of meaning explain that mystery and others.

[3] H. Woodrow and F. Lowell, Children's association frequency tables. Psychological Monographs, vol. 22, No. 97. A. J. Rosanoff, Manual of psychiatry (Sixth edition), pp. 546–620. Johnson O'Connor, Born that way.

Similarly the importance of the signs of "complexes" associated with certain words does not so much lie in the fact that the "complex" outweighs a normal interest in certain objects and facts, as in the fact that it outweighs a basal organization of habits, comparable to walking to the place where one wishes to be, swallowing what has been chewed, or holding a book right side up.

For the general study of interests and attitudes the free-association test with a word for a word is useful not so much because each response is in large measure the product of an individual interest, as because most of them are not. The material furnished is scanty, but it is presumably relatively free from efforts to make a good showing and from erroneous self-evaluations which the person has adopted. Its great merit as a test of interests is that the majority of the connections operate rather fluently as a consequence of the meaning-bonds and use-bonds so that the subject may be less on his guard in the few cases where his interests do determine responses.

I have examined the words written by fifty summer students in response to the words of the Kent-Rosanoff list (except 18 rough, omitted by inadvertence) read at the rate of one every five seconds, checking every response which is not simply and easily explained as the product of connections formed by frequency and fitness (1) in hearing and reading, (2) in speaking and writing, or (3) in the customary experiences of people in general. I have also checked responses like man, dress, fear, which may show individual interests greater in amount than the average if they occur often, in case they appear more than once in the hundred; also all failures to respond. The responses so checked are rather insignificant as revelations of the individual interests or experiences of the writers.

The numbers having 0, 1, 2, 3, and 4 responses checked were respectively 7, 10, 10, 4 and 4, or 70 per cent in all. Five had 6 checks; two had 11; there was one each for 7, 8, 9 and 12. There was one with 16; one with 19; one with 23; and one with 32.

The facts for each of those with 6 or more checks are reported below.

1. 6 man tall, eating butter, wish money, health purity, sheep ram, city of.
2. 6 short pie, woman hair, beautiful stings, girl skirts, doctor —, green —.
3. 6 hand person, wish letter, dream pleasant, boy suspenders, quiet study, afraid chasm.
4. 6 beautiful dress, yellow dress, blue dress, head comb, joy dancing, baby woman.
5. 6 comfort sick, slow wood, boy John, priest long, city dirt, blossom wrong.
6. 7 hand rings, short dress, woman mother, white dress, blue dress, moon love, afraid quiet.
7. 8 man crowd, butterfly monarch, spider afraid, dream mixed, boy away, religion hypocrite, moon scar, afraid night.
8. 9 music teacher, beautiful man, working man, soldier man, eagle brand, dream Freud, health drown, child man, doctor Jones.
9. 11 short man, woman man, citizen male, anger boy, girl boy, dream —, whiskey man, loud girl, thief —, tobacco —, scissors —, afraid boy.
10. 11 man dog, eating dead, comfort —, short —, bread —, boy bad, health Arizona, long division, child —, doctor bishop, baby Tommy.
11. 13 table to, dark of, comfort do, needle hen, trouble sin, dream ghost, red wine, memory Christ, head maul, butter ham, loud hate, bed slat.
12. 16 house —, cold —, rough —, high —, hard skin, lamp —, justice mule, boy —, health —, blue —, stove —, child —, tobacco fight, moon organ, street —, afraid load.
13. 19 table girl, chicken dark, sickness sing, command you, woman cat, river cow, working slovenly, trouble free, sheep duck, bath good, cottage river, head church,

whiskey school, doctor large, thief large, joy large, tobacco spit, moon bottle, blossom butter.

14. 23 soft hit, comfort vegetable, hand his, whistle Pineville, river drown, rough Gannisburg, citizen Davis, carpet with, dream Dixie, bread west, memory sweet, sheep clothes, cottage Hillsboro, ocean France, head Camp Harris, religion Henry, bitter what, thirsty righteous, doctor Arthur, joy to, baby sick, green sleep, afraid grape.

15. 32 man good, eating good, mutton red, comfort good, short arm, command boy, foot long, sleep long, anger long, girl good, earth large, eagle brave, dream long, boy nice, light strong, health good, bible large, hungry girl, priest man, ocean large, head large, religion good, whiskey old, hammer large, square city, doctor good, lion large, joy good, moon large, scissors good, king good, cheese butter.

The last four cases reveal a general dullness or disorganization or both more than special interests.

How far the revelations about these fourteen individuals would be indications of other tests of the same sort, and how far they would be substantiated by the real total facts of the individuals in question we need not now inquire. Nos. 4 and 6 may in reality be no more interested in personal adornment than Nos. 3 and 5. Nos. 8 and 9 may in reality be no more interested in sex than No. 10 is. The group included two clergymen, one of whom was the only one in the group to respond to *whiskey* by *good*. The other was the sole responder to *thirsty* by *beer*. This need not make us more suspicious of them than of the test. But the records do testify to the general naivete of the responses. And this is as essential to tests of interest and attitude as significance.

I have inspected 50 other similar responses to the Kent-Rosanoff list, all from women, inquiring in each case whether the responses seemed to indicate anything about the person's nature or history distinctive from other women in the group.

The indications are even fewer and less emphatic than in the fifty men and women reported above. One omitted responses to trouble, health, and bed, and may be suspected of having had some disease.

The writer of the following may be suspected of being in a healthy and optimistic mood: *house beautiful, citizen help, girl happy, sour apple, trouble smile, stomach food, health good, memory good, cottage lovely, hungry food, religion love, thief returned, baby lovely, moon beautiful, afraid not.*

One wrote *river geography, working problems, earth globe, boy school, health teacher, memory training,* and *thief child,* whence it would be reasonable to infer that she was a teacher. In general the responses seemed to differ little from those which a group of business women of equal intellect and education would give.

I have also counted the occurrences of *boy, man,* and *dress* in each of these 50 records and in records from 25 men of the same general group of summer-school students. There is no demonstrable difference; the frequencies (in per cents) being as follows:

	WOMEN	MEN
0	....	8
1	14	12
2	36	24
3	26	20
4	18	12
5	4	16
6	....	4
7	....	4
8	2	....

A test in which a subject is asked to say or write the first word he thinks of, or the first thing he thinks of, in response to a stimulus word is then usually not a test of what ideas and feelings the individual is accustomed to associate with the facts named by the stimulus word. The associations run more

directly from the stimulus word itself. One hears *man* and says *woman* or *boy* or *person* without delaying to think of the fact of manhood or of any one man and of what that fact or person suggests. If a foreign word like *oui* or *garçon* or *ja* is presented the response will be predominantly its English meaning, if that is known to the subject. If pictures should be inserted, the predominant response would be, I predict, to name them. If "afraid" is presented the response will be something which constitutes or aids the comprehension of the word fifteen times as often as something which the individual is afraid of. If he says "dark" that response is much more likely to mean that he learned the meaning of the word in connection with the dark than that he now fears the dark.

The test is probably not nearly so sensitive to the attitudes and interests of individuals as a test with sentences to complete by a word, or a test with choices after the pattern of the Strong tests of occupational interests. But it may well be revealing under certain conditions, and it will be more revealing now that the causation of ninety-five per cent of the responses in normal individuals is understood and can be allowed for.

ANALYTIC USE OF THE ARMY ALPHA EXAMINATION

DAVID WECHSLER

Psychiatric Division, Bellevue Hospital

The growing use of the Army Alpha for obtaining intelligence ratings on adults, both in clinics and in personnel work, has called attention to values of the tests other than that of mere classification based on the subject's total score. These have in part been met by translating total scores into mental age equivalents, but even more important would seem the need of analyzing the subject's performance on the individual tests which comprise the examination, in order to discover, as may be done through a detailed analysis of the Binet tests, any special abilities or disabilities. For this purpose the Army Alpha, indeed, lends itself even better than the Stanford-Binet because the same type of ability or at least material, is involved at all levels of performance. The question of course is how can this be done.

In the case of the Stanford-Binet psychologists now generally follow the suggestion made by Wells some years ago to consider an individual as showing special ability (or disability) along a given line when he scores a success in the same or similar type of test at least two years above (or below) his mental age rating. Thus if he attains a mental age rating of 8 years and passes the vocabulary test in year 10, he may be credited with having relatively superior language ability. The trouble is that only few tests on the Binet (the repetition of digits and the vocabulary being the conspicuous exceptions) recur throughout the entire scale. But it is precisely in this respect that Army Alpha enjoys a great advantage, for irrespective of the level (that is, total score) attained it is with the same type of material that the subject is tested. Furthermore, the individual tests on the Alpha measure more nearly what may be termed specific abilities than do a majority of

those which go to make up the Binet. Most obvious, of course, is Test 8 on the Army Alpha which measures range of information. Almost equally significant are Tests 2 and 6 which measure arithmetical reasoning ability. Then there are the common sense questions of Test 3 which may be taken as measures of comprehension of social relations. Test 4 (synonyms and antonyms) clearly correlates with language ability, and I am myself inclined to interpret Test 5 (the disarranged sentences) and Test 7 (the analogies) as measures of abstract reasoning ability.

Using the interpretations here suggested or any other which the examiner may find more appropriate, it now only remains to determine what score a subject must make on any of the individual tests in order that he may be rated as having a special aptitude or inferiority with respect to the ability which that particular test is interpreted as measuring. For this purpose I have had calculated from the data to be found in the Memoirs of the National Academy of Sciences the median scores on each of the eight individual tests[1] for various total scores. These are given in the accompanying table. Thus, if a subject makes a score of 63, one finds the interval 60 to 64, and running down this column one notes what median score the subject may be expected to make when his total score is within the limits indicated. In this instance it would be a median score of 7 on Test 3, a score of 5 on Test 7, and so on. As to the number of points which a subject must make in order that he may be rated as having special ability or deficiency along the lines measured by any particular test, I have found (empirically) that a difference of more than two points on tests 1, 2, 3, 5 and 7, and a difference of more than four points on tests 4 and 8 is significant. Subjects making scores on individual tests above the indicated medians for a particular total score by amounts greater than these may be said to have good or superior ability, and those making median scores smaller by similar amounts as having relatively poor ability in the particular functions measured.

Alpha Totals

Median Score on Separate Tests	Test	20–24	25–29	30–34	35–39	40–44	45–49	50–54	55–59	60–64	65–69	70–74	75–79	80–84	85–89	90–94	95–99	100–104	105–109	110–119	120–129	130–139	140–149	150–170
	1	3	3	4	5	5	5	5	5	5	6	6	6	6	7	7	8	8	9	9	9	9	9	11
	2	4	5	5	5	7	7	7	8	8	8	8	9	9	9	10	11	11	11	11	12	13	13	13
	3	2	3	3*	5	5	5	6	7	7	7	7	8	8	8	8	8	9	9	9	11	12	12	13
	4	1	1	1	1	3	5	5	7	9	10	11	12	13	15	15	17	17	19	21	23	25	29	31
	5	2	3	3	4	5	4	7	6	7	8	9	10	11	11	12	12	12	13	15	15	17	18	22
	6	2	4	4	4	4	5	5	6	7	8	7	8	8	9	9	9	10	11	10	11	11	12	12
	7	1	3	3	3	3	5	5	5	5	7	9	9	9	9	11	11	13	13	15	19	23	25	31
	8	5	7	7	9	11	11	13	13	13	15	17	17	19	21	21	22	23	23	27	29	29	33	35

* In a few instances the medians have been "corrected," that is altered (but never by more than one point), to conform to the trend of the separate test curves.

THE EFFECT OF TEMPORAL SET ON LEARNING

M. N. THISTED AND H. H. REMMERS
Purdue University

HISTORICAL

Aall (1) in 1913 tested two groups of Norwegian children for their recall of stories that had been told and groups of objects that had been shown to them. The one group was told that it would be tested at some later date while the other group was not told anything concerning a retention test. At the end of four weeks a test was given to both groups. The group with the time set scored from 4–13 per cent better on the stories and 18 per cent better on the objects that did the control group.

Peterson, (9) attempted to solve this problem: What difference will it make in later reproduction whether a person knows or does not know while reading a list of words that he will be asked to reproduce the words? He used 40 students of general psychology. Two lists of common words were used in the experiment. One list of words was read to the class and each student copied them on paper. They were told to turn the papers over and reproduce as many of the list as possible. These papers were then collected and a new list of words was given in the same way but with added warning that reproduction would be called for. Forty-eight hours later the students were asked to reproduce both lists of words. The results show that approximately 80 per cent of the students made better records with the determination of immediate recall than without it and 90 per cent made better records with determination in delayed reproduction. The gain in delayed recall is greater than the gain in immediate recall, being 50 per cent against 20 per cent. He concludes that it is natural to expect a greater gain with an active attitude.

About the time Peterson was working with a two-day time set, *Boswell and Foster* (3) were experimenting with the relative value of a 5-minute time set *versus* a two-week time set. Two series of 16 pairs of Chinese-English equivalents were given to four subjects, the one series being learned with the intent of remembering for 5 minutes and the other series being learned with a similar intent for a two-week period. Tests on the material were given at the end of the two periods. The results indicate that the longer time set is conducive to a more favorable memory attitude, but no positive conclusions were drawn because of the limited number of subjects. The difference was not statistically significant.

One of the pioneers in this field was Mulhall (7) whose learning material consisted of nonsense syllables and named photographs. One of the group of students arranged the nonsense syllables in four piles according to the ease of articulation and the photographs in piles according to their attractiveness, with no intention to recall. The experimental group went through a similar procedure but with the added instructions to "determine to remember." Both groups were tested immediately. As the author gives no criteria by which the two groups were equated his results suffer in validity. Subject to this criticism, the results indicate that learning with the intent to recall has a greater effect on reproducibility in proportion as the material is rich in associations.

Panicelli (8) formed two equated groups of 60 children each, one of which learned with intent to recall, while the other group had no intent to recall. The same story was read to both groups and both groups were tested on the recall of groups of nouns, verbs, etc. The experimental group reproduced a greater amount of parts of speech most frequently used by children and the words which were richest in association for them. Lacking statistical criteria, the results have little quantitative value, but indicate that voluntary effort in children (8–13 years) increases the recall of familiar words learned more or less incidentally in a context.

Within a larger study by Achilles (2) is found a smaller study in which he worked with photographs and nonsense syllables, testing two groups of subjects on recognition of names of photographs and the recall of syllables. He concludes that "determination to remember has little influence upon recognition, but considerable influence upon recall, and the influence of determination upon recall varies with the richness of associations of the experimental material."

The results of a study within a study by *Jones* (5) are unfavorable to the notion that mental set is of importance. He used three different classes under conditions which he considered prevented one class from knowing the type of experiment being conducted in two other classes. At the beginning of the class period slips of paper were handed out to all members of the three classes. The students seated on one side of the middle aisle received slips which contained the following instructions: "During the last five minutes of this period you will be given a written quiz on the material specifically discussed during the class period. As a part of the class will not be informed of the quiz, do not share this announcement with your neighbor. This is an experiment but it is important that you should make as high a rating as possible." The other half of the class received slips which contained reference material in the library. A test was given to all three classes and a comparison made of the average scores of the informed and uninformed groups. The difference between the means was only 1.5 ± 3.9. Jones' conclusions are as follows:

"The effect of expectation may be neglected as an important element influencing the average test scores of a class. The warned groups do not surpass the unwarned groups in either the immediate or the delayed tests and, under the conditions specified, a class will perform about equally well on a terminal quiz whether the quiz has been expected or not.—The primary factors determining the score—difficulty of material,

tempo, and other conditions of presentation—appear to operate with relative independence of the changes in mental set."

One of the most recent studies made on intention in learning is the one by *Sanderson* (*11*). His summary and conclusions follow verbatim: "An essential factor in learning is the attitude to learn. The daily routine of the ordinary individual exposes him to a multitude of unchanging sights and sounds, and yet he can give but a meager report of this constant stimulation. In this situation, the 'will to learn' is absent. It was the purpose of this experiment to investigate the effect of intent or mental attitude on the acquisition of two types of problems requiring some degree of motor ability. These problems were a stylus maze and a number sheet, the patterns of which were identical. The particular intention desired was instilled in the subject by special instructions. One phase of the general will-to learn in the mastery of a motor problem would appear to be the attention directed to the pattern itself. Accordingly, the intention selected for study was that mental set in which the subject should devote his attention to the learning of the pattern. ("You are to learn the pattern formed by these lines in such a way that I may call on you to reproduce it at the end of the learning.")

"What effect would it have on the learning of one or the other problem? What would be the subsequent effect on learning another similar problem without the intention? Control or speed groups and experimental or pattern groups composed of students of college level (128 males, 64 females) were compared. *Conclusions:* 1. It was found that the pattern intention retarded the speed of learning the maze and number-sheet. 2. But, when an individual was transferred from the number-sheet which had been learned under this intention, to learning a maze without the pattern intention, his time was lowered. 3. The effect of pattern intention was carried over to a repetition of the original problem. 4. The women showed less effect of the pattern instructions than the men."

It is evident that a condition of expectation of recall when injected into the initial instructions, has given variable and conflicting results. The types of material used were: stories, objects shown, lists of common words, Chinese-English equivalents, photographs, nonsense syllables, stylus maze, and number-sheets. A study by Jones minimizes the importance of mental set and he concludes that the factors in determining the score, which he calls the "primary" factors, such as difficulty of material, tempo, and other conditions of presentation, appear to operate independently of the mental set.

The results all told indicate that further experimentation is needed in which more subjects can be used and more typical learning situations made use of.

PROCEDURE

We shall present in the article a comparison of the results obtained when three different time-sets were used with 404 students in elementary psychology and educational psychology at Purdue University during the school years of 1929–30 and 1930–31. With few exceptions these students were of sophomore rank. Two groups, one experimental and one control, for each type of material and time-set were equated on the basis of sex and their average percentile scores on freshmen placement tests given a little more than a year previous to the present experiments.

The learning material consists of the following:

1. *Anglo-Saxon vocabulary,* in which fifty Anglo-Saxon words with their English translations were selected from the glossary of an Anglo-Saxon reader (4). No words were included which indicated obvious derivation. The following are examples of the pairs used: adl = disease; attor = poison; ea = river; etc. The students were given no information concerning the name of the vocabulary being learned.

2. *Religion of the Laplanders.* This article presents an instructive and interesting view of the religious creed and ceremonies of the Laplanders as practiced during the early part of

the 19th century (10). It contains much factual material which can be readily tested. As the volume from which the article was chosen is more than a century old, it was very improbable that any student would find the original article.

3. *Special Methods in Psychotherapy.* This article (12) presents in an interesting discussion some of the special methods utilized in psychotherapy today. It contains material of factual content considerably beyond the type of material with which the students had become familiar in a course in elementary psychology. As the author's name was not given, the name of the article abbreviated and the volume is to be found only on the reference shelf, it was assumed that no student would find the original article. The students testified that they had made no effort to obtain access to any of the materials used while the experiment was in progress. The time-sets consisted of an immediate, a three day, a one week and a two weeks' interval.

Anglo-Saxon material was used for two groups of 26 pairs, one group having an immediate time set and the other a three day time set. It was also used for two groups of 46 pairs in which the one group has an immediate time-set and the other a one week time-set. The mimeographed directions were as follows:

"On the lower half of this sheet you will find an exercise which will test your ability to learn new subject matter. A vigorous effort toward learning the exercise should give valuable results that may help provide an estimate of your general capacity for learning, and specifically, your ability to do independent work in this course. It has been found by experimentation that there is a close relationship between the ability to learn and intelligence.

"The exercise consists of fifty words with their English translations and you have twenty minutes in which to learn as many combinations as possible. Use your own method of study. *Do your very best.*

"You will be tested on this material immediately (for the two control groups) and three days from to-day or one week from to-day (for the two experimental groups)."

The religion of the Laplanders was the material used with two groups of 72 pairs, in which the control group had the immediate time-set and the experimental group had the one week set. The mimeographed directions were similar to those used with the Anglo-Saxon material, differing only in the reference to the type of material, the length of study period and the directions on time to be tested.

The article on Psychotherapy was mimeographed and given to two groups of 58 pairs, in which the control group had the immediate time-set and the experimental group the two weeks' time-set. The directions were similar to those given in the Anglo-Saxon experiment, differing only in reference to the type of material, the length of study period and the directions on the time of testing.

All students in the experiment took the immediate test given at the close of the study period. All students in the three-day-set study took the test given three days later. All students in the one week-set studies took the test given one week later. All students in the two week-set study took the test given two weeks later.

The tests of Anglo-Saxon material consisted of a haphazard arrangement of the Anglo-Saxon words in two columns of twenty-five words each, with the students instructed to fill in as many of the English equivalents as possible.

The tests in Religion of the Laplanders consisted of eighty objective type questions (42 true-false, 20 completion and 18 multiple-choice) constructed by one of the authors, and arranged in two forms, A and B.

The tests on Psychotherapy consisted of sixty objective type questions (40 true-false and 20 multiple choice) constructed by one of the authors and validated judgmentally by two members of the department of psychology. These were arranged in two forms, A and B.

The instructors conducting the classes included in the various studies received written directions to the effect that "at the end of the study period be sure to account for and collect all the mimeographed articles that were handed out." This was done to prevent the circulation of any copies of the learning material outside the class room.

As complete instructions were mimeographed and given to each student in all sections of this experiment, it is assumed that approximately the same interest-attitude towards the exercises was maintained in both experimental and control groups, the only significant variable being that of the temporal set or *Aufgabe*.

RESULTS

The experimental data are shown in Tables I to IV.

It is evident from Table I that the experimental group was somewhat at a disadvantage on immediate recall, but that, relatively, they gained on the control group in the delayed

TABLE I

Comparing the experimental and control groups on Anglo-Saxon material with a 3-day time-set

	IMMEDIATE RECALL			DELAYED RECALL					
	Av.	σ	$P.E._M$	Av.	σ	$P.E._M$	r	N	Av. gain or loss
Experimental	22.86	9.21	1.29	18.57	8.13	1.14	.860	26	−4.29
Control	26.55	8.46	1.19	19.74	7.62	1.07	.822	26	−6.81
Difference	−3.69 ± 1.75			−1.17 ± 1.56					
$\frac{\text{Diff.}}{\text{P.E.}_{\text{Diff.}}}$	2.11			− .75					

recall after three days. In neither immediate nor delayed recall is the difference statistically completely reliable.

When an interval of one week elapsed, however, a somewhat larger group of experimental subjects yielded a sig-

nificantly greater average performance for the experimental group with English and Anglo-Saxon equivalents (Table II).

TABLE II

Comparing the experimental and control groups on Anglo-Saxon material with a 1-week time-set

	IMMEDIATE RECALL			DELAYED RECALL					
	Av.	σ	$P.E._M$	Av.	σ	$P.E._M$	r	N	Av. gain or loss
Experimental	22.89	9.66	.96	15.66	7.71	.77	.860	46	− 7.23
Control	23.10	9.87	.98	10.62	6.72	.67	.822	46	−12.48
Difference	− .21 ±1.37			5.04 ±1.02					
$\frac{\text{Diff.}}{P.E._{Diff.}}$	.15			4.94					

The average difference is nearly a standard deviation of the raw score distribution on the delayed recall. The relative loss from forgetting is considerably less for the experimental group than it is for the control group.

With more meaningful material, Religion of the Laplanders, (Table III) and a week's temporal set, the results favor the

TABLE III

Comparing the experimental and control groups on "Religion of the Laplanders" material with a 1-week time-set

	IMMEDIATE RECALL			DELAYED RECALL					
	Av.	σ	$P.E._M$	Av.	σ	$P.E._M$	r	N	Av. gain or loss
Experimental	71.90	5.72	.45	72.18	5.38	.427	.856	72	+ .28
Control	72.14	6.18	.49	70.88	5.94	.47	.869	72	−1.26
Difference	−.24			1.30					
$\frac{\text{Diff.}}{P.E._{Diff.}}$	.36			2.16					

experimental group on delayed recall. There is a slight average gain for this group after the interval of one week. The average differences for both immediate and delayed recall are again statistically unreliable.

The material with which the experimental subjects were already somewhat familiar, Psychotherapy (Table IV) yielded

TABLE IV

Comparing the experimental and control groups on psychotherapy material with a 2-week time-set

	IMMEDIATE RECALL			DELAYED RECALL					
	Av.	σ	$P.E._M$	Av.	σ	$P.E._M$	r	N	Av. gain or loss
Experimental	38.79	8.45	.75	37.03	8.06	.71	.84	58	−1.76
Control	32.08	9.27	.82	30.88	9.00	.80	.67	58	−2.20
Difference	6.71 ±1.11			6.15 ±1.06					
$\frac{\text{Diff.}}{P.E._{\text{Diff.}}}$	6.05			5.80					

significant differential results on immediate as well as on delayed recall.

The tendency is towards a slower drop in the forgetting curve for the experimental group.

DISCUSSION AND CONCLUSIONS

The results shown in the foregoing tables were obtained under ordinary classroom conditions with such motivation as may be provided by this situation and the temporal set provided by the instructions. Theoretically one might expect a difference in favor of the experimental group in immediate recall. The reverse of this was found in all except one of the experiments (Table IV) here reported. In this last experiment there was also least evidence for a slower forgetting on

the part of the experimental group found in the other three experiments.

What would be true with more genuine motivation is a matter of speculation. There is evidence from previous experimental work for the statement that the ordinary classroom situation does not constitute a high degree of motivation (6). This variable, however, was presumably a constant for both groups. The differential effect of the *Aufgabe* has obvious implications for both theory and practice with reference to the effect of anticipated periodical tests and final examinations.

From the results of the present experiments under the conditions described, we venture to draw the following conclusions:

1. The experiments tend to establish a somewhat slower drop in the forgetting curve when a temporal set is introduced that delayed recall will be required.

2. The differences in average scores on delayed recall are not particularly impressive, either with reference to theoretical reliability or practical applications.

3. Experiments over a wider range of talent—grade school pupils, high school pupils, and college students—with actual school subject-matter as called for in semester examinations should be made.

BIBLIOGRAPHY

1. AALL, A. "Die Bedeutung der Zeitperspektive bei der Einprägung für die Dauer der Gedächtnisbilde." Ber. über d. r. Kongr. f. exper. Psychol., 1912, 237–47.
2. ACHILLES, E. M. "Experimental Studies in Recall and Recognition." Archiv. of Psychol., 1920, 6, No. 44.
3. BOSWELL, F. P., and FOSTER, W. S. "On Memorizing with the Intention of Permanently to Retain." J. of Psychol., 1916, 27, 420–460.
4. BRIGHT, JAMES W. Bright's Anglo-Saxon Reader. Holt, New York, 1913.
5. JONES, H. E. "Experimental Studies of College Teaching." Archiv. of Psychol., 10, No. 68.

6. KNIGHT, F. B., and REMMERS, H. H. "Fluctuations in Mental Production when Motivation is the Main Variable." J. of App. Psychol., Sept., 1932.
7. MULHALL, E. F. "Experimental Studies in Recall and Recognition." Amer. J. Psychol., 1915, 26, 217–28.
8. PANICELLI, E. "Influenza della Cosidetta Volonta di Apprendere sui Processi di Apprendimento." Rivista di Psicol., 1914, No. 2.
9. PETERSON, J. "The Effect of Attitude on Immediate and Delayed Reproduction." J. Ed. Psychol., 1916, 7, 523–32.
10. ROBBINS, THOMAS. "All Religions and Religious Ceremonies." 2nd Ed. Oliver Cooke & Sons, Hartford, 1824.
11. SANDERSON, S. "Intention in Motor-learning." J. Exp. Psychol. 1929, 12, 463–489.
12. TAYLOR, W. S. "Readings in Abnormal Psychology." D. Appleton and Co., New York, 1926, pp. 688–695.

TWO INDUSTRIAL APPLICATIONS OF A PSYCHOPHYSICAL METHOD

G. R. WENDT

Yale University, Institute of Human Relations

I. THE EFFECT OF SIZE OF PATTERN ON THE APPARENT BRIGHTNESS OF RUBBER TILE FLOORING[1]

This section describes a test of a method for the measurement of an apparent brightness difference between large and small patterns of the same type of rubber tile flooring.

On survey of a number of installations it had been noticed by the manufacturers that rubber tile floors did not appear identical in brightness when there was a difference in the size of the individual tiles making up the pattern. This difference had been very noticeable in two installations where the tiles had been laid in a checkerboard pattern with alternate light and dark squares. In one floor the individual tiles were six inches square, in the other nine inches square.

The problem therefore arose: how could this apparent brightness difference be measured? An adaptation of the Method of Constant Stimuli proved the most promising way. Two samples of flooring were made up, the units of one being twice the size of those in the other. These were simultaneously illuminated for short periods of time. One or the other always received a "standard" illumination. The other re-

[1] The work described in this paper was done in the Psychological Laboratory of Columbia University in the winter of 1929–1930. The investigation was sponsored by the Rubber Flooring Institute of the United States Rubber Company. The writer is indebted to Mr. A. G. McKinnon, Director of the United States Rubber Flooring Institute, for his suggestion of the problem and for his helpful cooperation.

The writer is very grateful to Professor A. T. Poffenberger, Professor R. S. Woodworth and the late Dr. C. Ladd-Franklin for time generously given for conferences and criticism.

ceived an illumination less than the standard by certain constant percentages. The subject indicated on each presentation whether the sample to his left or the sample to his right was the better illuminated. In this way the amount of the constant error was determined. The method had the advantage of approaching the conditions met in actual installations and of requiring no interpretation to apply the results.

APPARATUS

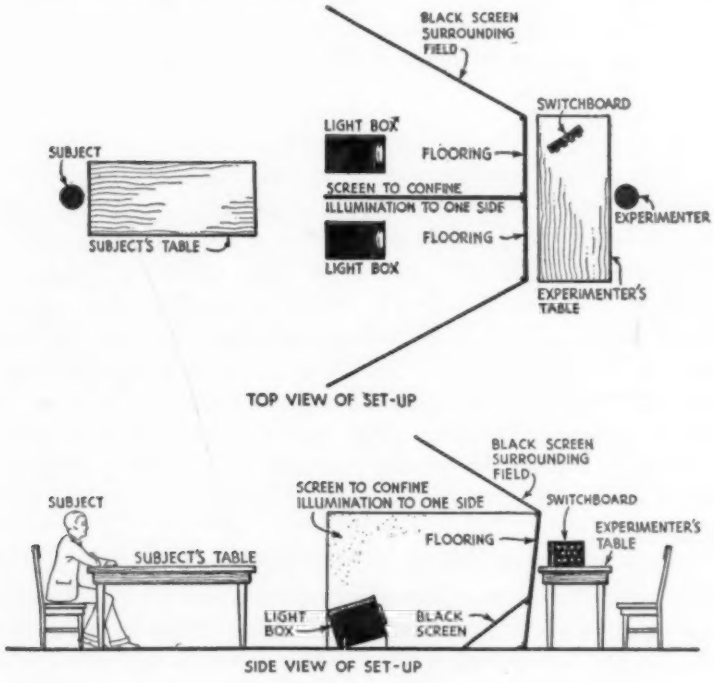
The flooring samples were two feet square. Each was a checkerboard pattern made up of alternate light and dark squares. In one the individual tiles were four inches square, in the other two inches square. The combination of tile used was one popular with those who buy the installations. The darker units had a dark background mottled with light material (total reflection factor of 17 per cent). The light tile reversed the relations, being a light background with a mottling of dark. Its reflection factor was 64 per cent.

These two samples were presented to the subject at a distance of fourteen feet. The two were placed side by side in a black-painted stand which separated them by 2½ inches. They stood upright so that their surface was at right angles to the line of regard. All surroundings in the field of vision were hung with uniform black cloth stretched to eliminate folds which might cause glare.

The illumination on each sample was independent from that on the other. A light box stood before each and a black screen placed between the two kept the illumination from spreading from side to side. The screen ran back eight feet from the center of the stand directly toward the subject. In this way the subject could see both samples simultaneously, but the screen, visible only in edge view, was not noticeable during the experimental series.

The illumination was generated by a bank of ten electric lamps (total of 325 watts) in each lamp box. It was trans-

mitted through opal glass, losing considerable intensity in transmission. They cast a "standard" illumination of about six foot-candles on the surface to be judged (as measured by an Edison foot-candle meter). The variations in the intensity of the illumination were obtained by cutting out one, two or three 25 watt bulbs from the illumination. This gave a per cent reduction of approximately 8 per cent, 15 per cent and 23 per cent respectively. Preliminary work had shown these variations to be suitable for such judgments.



PROCEDURE

Before any data were taken some preliminary work was done to find the best distribution of illumination variations. It was found necessary to give more presentations in which

the illumination on the small pattern exceeded that on the large pattern than presentations in which the large pattern received more. If this were not done the subject would make more judgments of "brighter" for the larger pattern and thus become aware of his constant error and make an effort to correct for it. The differences finally selected were as follows:

ILLUMINATION ON SMALL PATTERN	ILLUMINATION ON LARGE PATTERN
Per cent	Per cent
92	100
100	100
100	92
100	85
100	77

It was found necessary to give each subject several hours of preliminary training in the discrimination involved. This was done as follows: Two floor patterns were made up of Milton Bradley light and dark grey with reflection factors of about 80 per cent and 10 per cent respectively. The subject made judgments of illumination variations on these until he showed consistency. He was then asked to judge illumination differences on two patterns of rubber flooring without the factor of size difference. Both patterns were made of four-inch units. When he showed consistency on these, judgments were taken with two sizes of pattern (two-inch and four-inch squares).

Subjects in each case made fifty judgments in a series. A random order of ten presentations in which each difference appeared twice was presented five times. After each series a rest period intervened before the next series was run. The relative position of the large and small patterns was changed for each series to eliminate the space error and to cancel out apparatus inaccuracies.

The subjects were told that the illumination on the two patterns was being varied and were asked to give a rapid

judgment of which received the greater illumination. No equality judgments were allowed. None of the subjects knew the purpose of the experiment nor did they give any indication that they guessed.

SUBJECTS

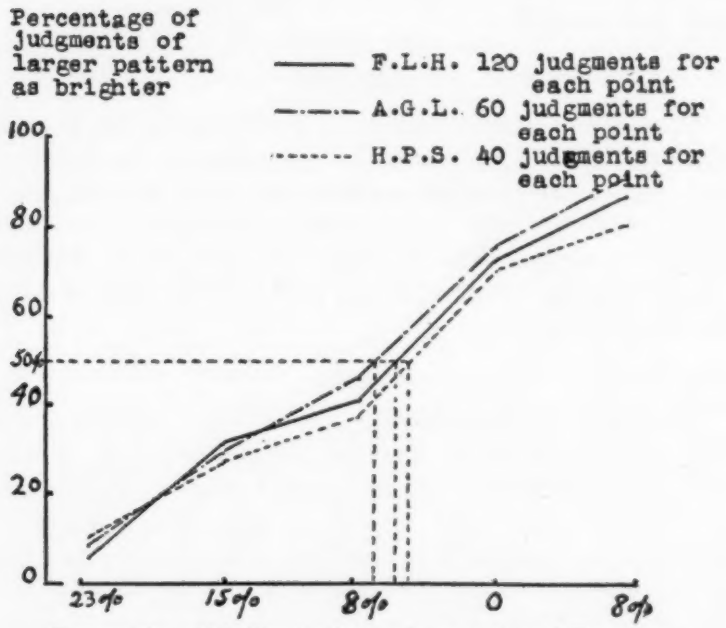
Three men graduate students in psychology served as paid subjects. Due to differences in time necessary for training and to the shortness of the available time the number of judgments made by each subject varies. One subject gave 120 judgments on *each difference,* the second gave 60 and the last 40. This gives a total of 220 on each difference and of 1100 for all differences.

RESULTS

The results are presented only graphically. In view of the lack of correspondence with the phi-gamma curve, it was deemed inadvisable to compute constant errors by the method of Fullerton and Cattell. By the graphic method the point of subjective equality is determined by the place where the 50 per cent line on the judgment axis cuts the curve of the psychometric function. This gives the constant error directly. The curves of each subject are shown on p. 274.

The results show that the greater apparent brightness of large patterns, as noticed in the actual installations, is in fact present. Under the conditions of this investigation the constant error for three subjects lies between 5 per cent and 7 per cent of the illumination used. That is, it required five to seven per cent more illumination on the small pattern to make it appear equally illuminated with the large. In view of the peculiar shapes of the curves found it may be that these values should be larger. It should be pointed out that the curve of each subject shows the same deviations from the usual ogive curve.

CURVES OF EACH SUBJECT



Per cent by which illumination on large pattern differs from illumination on small pattern.

There was no indication that the constant error increased or decreased as a result of practice in making the discrimination.

DISCUSSION

These results are obviously limited by their preliminary nature. The most important limitations seem to be the following: (1) To duplicate the conditions found in installations the flooring samples should be larger than those used (two feet square) and should be placed in the horizontal position. (2) Good technique demands the elimination of the use of a

dark room and dark surroundings and the substitution of light surroundings and a light-adapted subject. (3) The results cannot be applied to other relative or absolute sizes. Contrast and color constancy phenomena are known to vary with the square root of the area, but since the cause of the present phenomenon is not known, the effect of the size factor must be experimentally determined. (4) The results cannot be applied to other standard illuminations. Color constancy phenomena are known to be more marked under reduced illumination (Katz, D. Die Erscheinungsweisen der Farben und ihre Beeinflussung durch die individuelle Erfahrung. *Zsch. f. Psychol. u. Physiol. d. Sinnesorg.* 7. Ergbd. 1911, pp. xviii + 425) but whether this would be the case here cannot be stated before experiment. (5) The cause of the constant error is not apparent from the results. Explanations in terms of contrast, irradiation, "constancy" effects and habitual association of brightness with size have been suggested. Experiments designed to test each hypothesis could be devised and would be of great help in the application of the results.

II. THE EFFECT OF COVER MARKINGS ON THE APPARENT SIZE OF GOLF BALLS

When, a short time ago, the size of the regulation golf ball was increased by 0.06 inch diameter, a well-known golfer was observed to be using a recess-marked ball (small round cup-like depressions on the surface). Since the recess-marked ball is generally known to achieve slightly less distance than the mesh-marked ball (angular indentations), he was asked the reason for his preference. It developed that he felt the recess-marked ball to appear smaller and hence preferred it for the added confidence it gave on the putting green. Out of this circumstance a problem arose.

To test for the presence of such an illusion of size-difference the large (1.68 inches) and small (1.62 inches) balls in mesh and recess markings were presented repeatedly in the twelve

possible pairs. The balls were placed on a close-clipped lawn at 5½ feet distance from the eyes and exposed for one second.

Two subjects only were used. Neither showed the presence of the illusion sought for.

The illusion noticed by the golfer was probably one of the class best called "individual" illusions.

SUMMARY

The Method of Constant Stimuli was applied to two problems in the measurement of psychological factors in industrial products. Using this method it was shown that a rubber tile floor composed of alternate light and dark squares will appear brighter when its pattern is made up of larger units. The amount of this apparent brightness difference was measured. In a second application of the same method it was found that the character of cover markings of golf balls has no effect on their apparent size.

Such applications of experimental psychological methods seem to the writer to offer a wide and as yet untouched field for work in applied psychology. The applications of psychology to industry have included personnel studies, studies relating to the adjustment of the worker to his job and studies in the psychology of advertising. Except for an occasional study of consumers' preferences little measurement of psychological factors in industrial products has been attempted. The field is one which recommends itself not only because of its range of possibilities but also because of the relative ease with which its problems can be solved.

HABITUAL WORD ASSOCIATIONS TO COLORS AS A POSSIBLE FACTOR IN ADVERTISING

ROY M. DORCUS

Psychological Laboratory, Johns Hopkins University

In the field of advertising there is a definite attempt to establish associative bonds between the article advertised and the particular color used in presenting the article to the public. The general theory upon which this procedure is based is that when a given color is seen or thought of again, the article advertised will be immediately associated with it. This form of backward association has not been adequately investigated. The author feels that an investigation bearing on this point would serve to clarify certain features of color association. The specific questions which this investigation attempts to answer may be stated briefly as follows:

(1) Does an individual tend to associate the same color with the word which has been associated previously when that color was presented?

(2) What will be the individual's color response to the words most frequently given by a group in response to the colors presented?

(3) What will be the tendency concerning the color responses to a word which has been associated with more than one color?

In this experiment 131 male students from the Johns Hopkins University and the University of Maryland were used as subjects. The colors presented were red, orange, yellow, green, blue and purple. The saturation and brightness of each color was the same as reported in "Color Preferences and Color Associations."[1] For each color shown the subjects

[1] For the exact saturation and brightness of the colors see author's article entitled "Color Preferences and Color Associations." Ped. Sem. and J. of Genetic Psychol., 1926, 33, 399–434.

were required to write down the first word they thought of other than the name of the color itself. When this was finished the colors were again presented and they were asked to write down the name of each color. About eight weeks later each individual who had participated in the first part of the experiment was given a little booklet containing the twelve words which he had responded to the colors. These were arranged one to a sheet so that by turning over the sheets the one seen was removed and the next one exposed. They were asked to write on these sheets the name of the color which the word on the sheet suggested. At the same session a list of 51 words which had been selected on a frequency basis from all the words associated with the colors on the former presentation was read and the color responses to these words were obtained.

Table 1, column 1, shows a list of the original colors presented. Column 2 shows the percentage of individuals who gave in response to their own words the original color and also the percentage (over 10 per cent) of individuals who gave in response to their own words a color other than the original color shown. An examination of the table shows that the association from the color to the word and from the word back to the same color is fairly uniform in the case of red, blue, and green. Sixty-one per cent of the individuals respond yellow to the words associated with yellow and 49.8 per cent respond purple to words associated with purple. To the words associated with the color orange, yellow is as likely to be the response as is orange. The actual percentages are 37.9 and 39.0 for yellow and orange respectively.

In the unsaturated series a color other than the original color shown has an equal or greater chance of being associated with the words except in the case of blue and green. The original unsaturated yellow is replaced by green and brown and the original unsaturated red is replaced by pink, brown, and white. The pink, however, is very similar in color to the unsaturated red.

TABLE 1*

ORIGINAL COLOR PRESENTED	INDIVIDUALS' COLOR RESPONSES TO THEIR OWN WORDS	
Saturated series	Color	Percentage
Red	Red	82.1
Orange	Yellow	37.9
	Orange	39.0
Yellow	Yellow	61.5
Green	Green	77.7
Blue	Blue	79.9
Purple	Purple	49.8
Unsaturated series		
Yellow	Green	23.3
	Brown	18.4
	Yellow	12.3
Green	Green	44.3
	Blue	14.3
Red	Red	22.8
	Pink	21.4
	Brown	10.0
	White	10.0
Orange	Brown	42.2
	Yellow	10.5
Blue	Blue	67.8
Purple	Purple	32.3
	Blue	21.5

* The last column contains only those colors which were given in response by more than 10 per cent of the subjects.

Table 2 shows the most frequent words and the percentage of individuals giving that word for each color. It also shows the percentage of individuals (over 10 per cent) who responded with any given color to each word. For example, 9.7 per cent associated apple with the color red. When the word apple was the stimulus word 82.6 per cent of the responses were red and 13.2 per cent of the responses were

TABLE 2

ORIGINAL COLOR PRESENTED	WORDS FREQUENTLY ASSOCIATED WITH THE COLOR		COLOR RESPONSES TO WORDS	
	Words	Percentage	Color	Percentage
Red	Apple	9.7	Red	82.6
			Green	13.2
Unsat. yellow	Army	2.7	Khaki	44.7
			Brown	29.3
Unsat. red	Baby	3.8	Pink	47.3
Unsat. blue		2.5	White	33.3
			Blue	10.7
Yellow	Banana	3.1	Yellow	97.9
Orange	Black	2.4	Black	62.2
			White	21.1
Red	Blood	6.7	Red	93.9
Red	Blue	2.2	Blue	64.3
			Green	13.7
Unsat. red	Brick	4.8	Red	85.7
Unsat. yellow	Buff	5.4	Gray	50.0
			Brown	17.0
Unsat. green	Ceiling	3.1	White	64.5
			Yellow	14.5
Unsat. orange	Chalk	3.1	White	95.9
Unsat. orange	Clay	3.1	Red	30.8
			Yellow	25.5
			Brown	17.0
			Gray	15.9
Purple	Cloth	2.4	White	41.7
			Blue	15.3
			Black	12.0
Yellow	Coward	2.3	Yellow	89.6
Yellow	Crayon	2.3	Blue	26.3
			Red	25.2
			Green	10.5
Red	Danger	7.5	Red	96.9
Unsat. yellow	Dirt	6.3	Black	44.8
Unsat. orange		3.1	Brown	41.8
			Gray	10.2
Unsat. yellow	Drab	2.7	Gray	29.4
			Olive	21.1
			Brown	11.7

TABLE 2—(*Continued*)

ORIGINAL COLOR PRESENTED	WORDS FREQUENTLY ASSOCIATED WITH THE COLOR		COLOR RESPONSES TO WORDS	
	Words	Percentage	Color	Percentage
Red		8.2		
Yellow		2.3		
Blue		6.1	Blue	27.9
Purple		11.2	Pink	13.9
Unsat. green	Dress	7.2	Red	13.9
Unsat. red		11.5	White	13.9
Unsat. blue		6.7		
Unsat. purple		14.2		
Unsat. orange	Earth	3.0	Brown	58.7
			Black	19.5
			Gray	10.3
Orange	Egg	2.4	White	47.4
			Yellow	39.3
Red	Fire	9.7	Red	80.6
Red	Flag	3.0	Red	61.7
		2.3	Blue	12.7
Yellow		3.1	Red	26.8
Purple	Flower	5.6	Pink	19.3
Unsat. purple		3.0	White	18.2
			Yellow	11.8
Orange	Fruit	14.8	Yellow	31.9
			Green	21.2
			Red	20.1
			Orange	14.8
Unsat. green	Grass	6.2	Green	96.9
Green		41.9		
Purple	Ink	2.4	Blue	52.5
			Black	36.3
Green	Irish	3.8	Green	91.9
Green	Ireland	3.0	Green	88.1
Purple	King	3.2	Purple	51.7
			Red	13.1
			Gold	12.0
Unsat. orange	Leaf	3.1	Green	78.7
			Brown	18.1
Yellow	Lemon	25.1	Yellow	92.9
Unsat. yellow		4.5		
Unsat. yellow	Mud	2.7	Brown	51.0
Unsat. orange		6.1	Black	26.5
			Yellow	10.2

TABLE 2—(*Continued*)

ORIGINAL COLOR PRESENTED	WORDS FREQUENTLY ASSOCIATED WITH THE COLOR		COLOR RESPONSES TO WORDS	
	Words	Percentage	Color	Percentage
Unsat. yellow	Mustard	9.9	Yellow	51.5
			Brown	26.7
Unsat. yellow	Olive	13.5	Green	77.7
			Olive	12.1
Orange	Oranges	18.1	Orange	78.7
Yellow		3.1	Yellow	17.1
Yellow	Paint	2.3	Red	26.9
Unsat. yellow		3.6	White	25.8
			Green	11.8
			Yellow	10.7
Red	Paper	2.2	White	93.8
Red	Passion	2.2	Red	76.3
Purple		3.2		
Green	Peas	2.2	Green	94.8
Unsat. green		7.2		
Purple	Plum	2.4	Purple	61.4
			Blue	15.6
			Red	12.4
Unsat. green	Robin egg	3.1	Blue	66.6
			White	10.7
Unsat. red	Rose	4.8	Red	64.2
			Pink	20.4
Purple	Royalty	6.4	Purple	51.0
			Red	20.6
Blue	Sea	6.1	Blue	59.1
		30.5	Green	34.6
Blue	Sky	46.9	Blue	90.9
Orange	Sun	2.4	Yellow	15.9
Yellow		2.3	Red	27.5
			White	11.2
Orange	Sunset	3.3	Red	49.4
			Orange	20.6
Green	Tree	5.3	Green	57.5
			Brown	36.3
Purple	Violet	4.0	Purple	42.8
			Violet	41.8
Blue	Water	2.3	Blue	35.6
Unsat. blue		3.3	White	21.8
			Green	11.5
			Gray	10.3

green. Dress was associated with practically all the colors shown, yet when the word dress was presented, blue, pink, red, and white were the colors given. When the color green was presented 41.9 per cent of the individuals associated grass with it; when grass was presented 96.9 per cent of the individuals gave in response green. The above relation between the percentage of individuals giving the word and the percentage of individuals giving the color holds true for the color blue and the word sky, also for the color yellow and the word lemon. For the color yellow and the word banana a slightly different relation exists. Although only 3.1 per cent of the subjects associated banana with the color yellow, when the word banana was the stimulus 97.9 per cent respond with the color yellow. A similar relation holds for red—blood, green—Irish or Ireland, red—danger, and red—brick. Certain other words and colors show a quite different associative relation. For example, the association is from unsaturated green to ceiling to white, from unsaturated orange to chalk to white, from unsaturated orange to leaf to green, and from red to paper to white. Each word just mentioned was associated with its given color by only a small percentage of the subjects; when that word was presented the majority of the subjects associated it with some color other than the one originally shown.

The color responses to each individual's own words and the color responses to the words given most frequently by the group can be compared in tables 1 and 2. For instance, 82.1 per cent of the individuals responded red to the words which they gave in response when the original color shown was red. When the word apple was presented 82.6 per cent of the responses were red although only 9.7 per cent of the subjects had associated apple with that color. Danger was associated with the saturated red by 7.5 per cent of the individuals and in response to danger 96.9 per cent gave the color red. It is also found that the colors which the individuals tend to associate with the words they gave in response to the unsaturated

red are red, pink, brown, and white. The words associated most frequently with the unsaturated red are baby (3.8 per cent), brick (4.8 per cent), dress (11.5 per cent), and rose (4.8 per cent). Rose and brick give a higher percentage of red responses than do the words of each individual; baby and dress give about the same percentage of red responses as are given by the individuals to the words that they associated with the unsaturated red. From the above comparison of the data of tables 1 and 2, it appears then that in the case of saturated colors about the same percentage of individuals that respond with the original colors to the words they have already associated with that color, will also, when presented with the words most frequently associated by the group with given colors, react to each word with the color which originally furnished the stimulus for that word. There is a tendency in the unsaturated series for this relation to be set aside.

Table 3 shows the frequency of the color names assigned to the colors presented. It is evident that there was but little disagreement in assigning names to the saturated red, the saturated yellow, the saturated green, and the saturated blue. Orange and purple in the same series occasioned more difficulty. Orange was called yellow by 17.1 per cent of the subjects. Purple was called violet by 11.2 per cent and lavender by 6.8 per cent.

In the unsaturated series green was called green by 90.6 per cent of the individuals and blue was called blue by 94.7 per cent of the individuals. Red was called pink by 79.4 per cent, orange was called brown by 78 per cent, and the other two colors were called a variety of names. There was little agreement in naming the purple and yellow. This difficulty in naming the unsaturated colors may be responsible for the wide range of color responses to certain of the words. It is also quite evident that some of the subjects employed certain color names with which they were not adequately acquainted.

It may be concluded then:

TABLE 3

ORIGINAL COLOR PRESENTED	NAMES ASSIGNED TO COLORS	PERCENTAGE
Red	Red	94.0
	Crimson	2.2
	Vermillion	2.2
	Carmine	.7
	Dark pink	.7
Orange	Orange	79.8
	Yellow	17.1
	Banana	.7
	Buff	.7
	Peach	.7
	Brown	.7
Yellow	Yellow	95.6
	Orange	2.2
	Green	2.2
Green	Green	97.8
	Blue	1.5
	Red	.7
Blue	Blue	98.6
	Purple	.7
	Brown	.7
Purple	Purple	79.9
	Violet	11.2
	Lavender	6.8
	Brown	.7
	Gray-red	.7
	Blue	.7
Unsat. yellow	Yellow	45.7
	Green	19.0
	Olive	16.4
	Drab	5.2
	Buff	4.3
	Brown	2.6
	Lemon	1.8
	Tan	1.8
	Gray	.8
	Henna	.8
	Mustard	.8
	Purple	.8
Unsat. green	Green	90.6
	Blue	7.1
	Pink	2.3

TABLE 3—(*Continued*)

ORIGINAL COLOR PRESENTED	NAMES ASSIGNED TO COLORS	PERCENTAGE
Unsat. red	Pink	79.4
	Salmon	4.1
	Terra cotta	4.1
	Brick	4.1
	Brown	3.3
	Orange	3.3
	Vermillion	.8
	Buff	.8
Unsat. orange	Brown	78.0
	Tan	4.0
	Yellow	2.4
	Blue	2.4
	Buff	1.6
	Terra cotta	1.6
	Cream	.8
	Drab	.8
	Dun	.8
	Flesh	.8
	Chocolate	.8
	Cordovan	.8
	Ochre	.8
	Red	.8
	Putty	.8
	Maroon	.8
	Orange-red	.8
	Rustic	.8
Unsat. blue	Blue	94.7
	Violet	2.3
	Green	1.5
	Brown	1.5
Unsat. purple	Purple	35.6
	Lavender	30.3
	Violet	17.9
	Blue	5.3
	Lilac	4.7
	Orchid	1.5
	Pink	1.5
	Heliotrope	.8
	Magenta	.8
	Old rose	.8
	Scarlet	.8

(1) That the individual does tend to associate the same color with the word which has been associated previously with

that color. This is especially true when the colors are readily named, such as is the case with saturated colors. Of the saturated colors blue, green, and red are given in response to words previously associated with those colors by a large majority of the subjects. In the unsaturated series blue is the only color for which the above condition holds true.

(2) That there is very little relation between the number of times a given word has been associated with a given color and the number of times that that color will be given in response to that particular word. For example, Irish was associated with green by 3.8 per cent of the individuals and green was later associated with Irish by 91.9 per cent of the individuals. Grass was also associated with green by 41.9 per cent of the individuals and green was in turn associated with grass by 96.9 per cent of the individuals.

(3) That a word that has been associated with more than one color will not have associated predominantly with it any one color. The word *dress* serves as a good illustration of this point.

(4) That there is rather general agreement in naming the saturated colors and that there is considerable disagreement in naming the unsaturated colors except blue and green.

(5) That this method may be applied successfully to specific kinds of merchandise, thereby furnishing information concerning habitual associations to specific articles aside from the colors usually employed to advertise those articles. This is especially pertinent to colors that are difficult to name.

A TECHNIQUE FOR DETERMINING THE RELATIONSHIP BETWEEN REPORTED PREFERENCE AND PERFORMANCE[1]

HERBERT LLOYD BOWMAN
Bowling Green College of Commerce

INTRODUCTION

The theory that through interest success in a learning activity is more probable is generally accepted by modern educators. The acceptance of this theory is reflected in the practice of utilizing children's interests or preferences in the determination of both subject-matter and schoolroom methods. Likewise, the so-called interests of the individual serve as bases in giving educational and vocational guidance, the assumption being, apparently, that one is more likely to succeed in an activity for which an interest or liking is expressed. Such an assumption gives rise to the question of whether or not a mere indication of interest in an activity is symptomatic of a specific native ability in that activity.

A study of educational literature reveals that a number of investigations has been made during recent years for the purpose of determining what relationship exists between interest and ability. For the most part these studies have been of two types: (1) Those concerning the relation of interest to ability in various school subjects or broad fields of subject-matter, and (2) those dealing with interest and ability in vocational occupations and activities. Conclusions reached in these studies are not in agreement as to the possibility of predicting ability in a field of activities from manifested in-

[1] For complete account of study see: Bowman, Herbert Lloyd. *The Relation of Reported Preference to Performance in Problem Solving.* University of Missouri Bulletin, Vol. 30, No. 36. Columbia: University of Missouri. 1929. 52 p.

terest in it, neither do they agree with respect to the stability of early interests.

The difficulty in determining this relationship arises, in part, from the inadequacy of the measuring instruments employed in objectively recording the interest reactions. The sources thus obtained lend themselves with difficulty to reliable analyses and conclusions. The inventory method has been used, as has also the rating scale and the questionnaire. A further complication has arisen due to the fact that the relation of interest to ability has been studied principally from the standpoint of broad fields of subject-matter. Preferences for certain school subjects or for certain vocational activities have been obtained and compared with achievement in these subjects or fields. This condition favors the entrance of many factors into the measurement of the two variables, thus increasing the subjectivity and consequently the difficulty of accurate measurement and determination.

In a recent study the writer undertook to obtain the relationship between interest and ability in a rather narrow field, namely, problem solving in arithmetic by junior high-school pupils. The aim of this article is to describe the method used in obtaining interest scores, and in arriving at a coefficient of correlation between the two variables.

MEANING OF INTEREST

In using the term interest we refer to an individual's attitude, and his attitude is his mode of reacting to a situation. Attitudes are not necessarily separate reactions in themselves, but are aspects of the individual's larger reactions to situations. Attitudes are primarily of two types, aggressive or positive on the one hand, and retreating or negative on the other. In intellectual situations the positive reactions are called interest.

Interest is almost always comparative. We are more stimulated by one thing than by another and consequently react more vigorously in a positive manner toward the one than

toward the other. There is no way of measuring the depth or extent of interest. We can only get a report of preference of one situation over another when two or more situations are presented. The preference may then be evaluated by assigning to it a score based on some system of weighting, such scores meaning just what the weighting means.

As stated above we endeavored to determine the relation between reported preference and performance in problem solving. The expression reported preference rather than the term interest was employed, and it means that, of a group of problems presented to a pupil to be solved, one problem is reported as being better liked or more preferred than any others of the group.

PROCEDURE

The nature of the problem made it necessary to employ a test which would furnish objective measurements of the two variables, preference and performance. In undertaking the construction of an adequate instrument it was obviously necessary to provide for two things. First, arithmetic problems had to be presented to the pupils for actual solution, and second, provision had to be made for obtaining reports of preference for certain of these problems over others. Furthermore, it was necessary that the problems represent certain definite types throughout the test in order to provide for differentiation and therefore a basis for comparison.

Fifty problems were selected for use and these were made up into two test forms of twenty-five problems each. The differentiated types of problems used were five in number, and were as follows: (1) Problems based upon adult life activities; (2) problems based upon child life activities; (3) problems whose descriptions or settings were in the field of nature or science; (4) problems so stated as to take on the nature of a puzzle; and (5) problems of pure computation only, in which the directions for the right procedure were given. Each form of the test was made up of five pages with five

problems to a page. One of each of the five types occurred on each page, and the positions of the types were rotated so that no type occupied the same position on any two pages.

The pupils were instructed to solve the problems, then to fill out an incomplete sentence found at the bottom of the page, which read, "The problem on this page I like best is No. ——." In order that the preference for a problem be based on the problem situation or setting and not upon the element of difficulty of computation, it was necessary to hold constant this difficulty factor among the problems of each group or page. It is recognized that a number of factors other than that of computation affect the difficulty of solution of a problem. Among these are objects named, the presence of technical and unfamiliar words, the general setting, etc. No effort was made, however, to equalize the difficulty of the problems with respect to these factors, since they were the very elements which differentiated the problems into types.

Following a try-out of the test and a consequent slight revision, the forms were administered to 564 pupils of the 7th, 8th, and 9B grades. The performance reliability of the whole test was determined by correlating the results of Form I with those of Form II. This gave a result of: $r = .90 \pm .005$. Treating all fifty problems as one test and calculating the performance reliability of fifty against fifty reactions, it was found that $r = .95 \pm .003$. This represented the true performance reliability. Pursuing the same procedure with preference scores, the derivation of which is described in the next section, we obtained: $r = .77 \pm .014$. In this case we were correlating ten reactions with another ten. Such a limited range is always productive of a low r. Using the Spearman-Brown formula so as to raise the preference reactions to fifty items, we obtained an r of .94. Thus the coefficients of correlation indicate a considerable degree of consistency in measuring the two variables.

The question of the validity of the preference reports arises

at this point. We wish to make clear the point that we were concerned with *reported* preference. As previously indicated, no method has yet been devised that is known to measure objectively an individual's interests or likes as he feels them, if indeed that means anything. We can obtain only the oral or written reports of what the individual *says* that he likes. There is, consequently, no criterion by which the preference scores could be compared for validation. We know only that the individuals reported preferences, and did so with remarkable consistency. We have treated those scores for what they are, namely, *reported preferences.*

The validity of the research test with respect to performance was determined by correlating the performance scores with certain well known criteria. Table I shows the result in this connection.

TABLE I
Correlations of research test with certain criteria

	RESEARCH TEST
Stanford Arithmetic Reasoning Test	.82 ± .01
Stevenson's Arithmetic Reading Test	.72 ± .01
Otis Self-Administering Test of Mental Ability	.63 ± .02
Teachers' Marks	.60 ± .02

SCORING THE TEST PERFORMANCE

Each of the two test forms was scored for the number of problems correctly solved. The forms were then combined and treated as one test, thus providing for a total possible performance score of fifty. No analysis was made of the procedure used by the pupils in solving the problems, and no partial credit was given for the solutions. The problem was merely marked as right and a credit of one given if the correct answer was obtained, and wrong if the answer was incorrect or if no answer was given. A distribution of the performance scores are not shown here, but it was fairly symmetrical with a mean performance of 21.99 P. E. Mean 0.35.

PREFERENCE SCORES

As has been stated, the twenty-five problems of each test form were arranged in five groups of five problems to a group. The subject was asked to indicate, after endeavoring to solve the five problems of a group, which one of the five he liked best. The differentiation between the problems of one group was based upon the situations or settings which surrounded or clothed the actual problem. As has been explained, these settings were of five types, namely, (1) the adult type, (2) the child type, (3) the science type, (4) the puzzle type, and (5) the pure computation type. In treating the results we dealt with individual and total preference as preferences for *types* of problems, and not as preferences for particular or specific problems of the fifty included in the test.

In order to obtain an arithmetical value for each preference, and consequently a total preference score for each pupil, it was necessary to assign to each of the problem types some numerical value. This value or weight was determined by ranking the types according to the total number of preferences reported by all the pupils for each type of problem. For example, the type for which the greatest number of preferences was reported was ranked first and assigned the value of one; the type for which the second greatest number of preferences was reported was given the value of two, and so on through the five types. An analysis of the reported preferences with respect to types revealed that the computation type ranked first in number of preferences received, the child type second, the adult type third, the science type fourth, and the puzzle type fifth. Based upon this order of ranking, then, we assigned weights or values to the types as follows:

computation type	1
child type	2
adult type	3
science type	4
puzzle type	5

The next step was actually to determine the preference score for each pupil. This was done as follows: The number of preferences for each type of problem was obtained for the individual pupil. These numbers were then multiplied, respectively, by the value of the type which they represented. The following example will serve to illustrate the method. Case 13 of the 8th grade expressed three preferences for the adult type, five for the child type, and two for the science type. His preference score was determined thus:

3 times 3 = 9
5 times 2 = 10
2 times 4 = 8

Score = 27

In order to insure entire comparability of preference scores, only those of the individuals who reported ten, the maximum number of preferences possible, were determined and used in the calculations. There were 369 pupils who made complete reports of preference. The preference scores for these pupils are shown in Table II.

An inspection of the table containing the preference scores shows that the distribution is fairly symmetrical, the mean being 25.41 and the median 25.80. Reverting to the description of how the preference scores were derived, it will be remembered that the pure computation type of problem had a preference value of one. It can be seen, then, that the lowest possible score was ten, for an individual making this type his choice in all instances would have a score of ten times one, or ten. On the other hand, the maximum score possible occurred when an individual consistently preferred the fifth ranking type. In such an instance his score was fifty.

The question may arise as to the meaning of a high or low preference score. It will be recalled that the weight given to each type was determined by the way the pupils as a group ranked them, the type most often preferred being given the

TABLE II
Distribution of 369 cases with respect to preference scores

SCORES	F
40–42	3
38–40	4
36–38	8
34–36	8
32–34	26
30–32	45
28–30	40
26–28	46
24–26	45
22–24	35
20–22	30
18–20	35
16–18	15
14–16	11
12–14	11
10–12	7
Median	24.88
Mean	25.20
S. D.	6.26
P. E. Mean	0.22

value of one, etc. A low preference score would mean, then, that the pupil making it reported preferences that were somewhat in agreement with what the group preferred. On the other hand, the pupil who scored high in preference necessarily chose in general as his preference the types which ranked lower in number of preferences by the whole group. Had we ranked the preferences in a reverse order the results would have appeared different but would have meant precisely what the scores herein reported stand for.

CORRELATION BETWEEN REPORTED PREFERENCE AND PERFORMANCE

The original purpose of our study was to determine the degree and significance of the relationship between reported

preference and performance with respect to arithmetic problem solving by pupils of the junior high-school level. A comparison of the two sets of scores by means of coefficients of correlation revealed something of the degree of correspondence between the two variables, and these coefficients will be indicated at this point.

When the derived preference scores were correlated with the total performance scores a coefficient of correlation of $.46 \pm .03$ was obtaind. Since the preference scores were based upon the rank given by the group to the types of problems, we likewise, and by a similar method, obtained derived performance scores by weighting the score on each type according to that type's rank in performance by the whole group. By correlating these two sets of derived scores a correlation of $.58 \pm .02$ was discovered.

The above two coefficients of correlation were obtained when preference was correlated with *total* performance, that is, performance on all fifty problems. We were interested to learn what correlation existed between preference and performance on *only* the ten problems preferred. Accordingly these ten problems that were preferred by each pupil were scored for correct solution, then each one solved correctly was weighted according to the rank in performance given to the types by the group. When these two sets of derived scores were correlated the result was a coefficient of correlation of $.56 \pm .02$. This is probably the best expression of the relation existing between reported preference and performance. It means that when unusualness of preference is given high value and unusualness of performance is valued in the same way, the correlation between preference and performance on the items preferred is as indicated above.

The coefficients of correlation just reported are not high enough to be of predictive value, yet they are markedly positive and considerable significance may be attached to them. They mean that there is positive correspondence between high performance in arithmetic and unusualness of preference, the

individual scoring high in performance on the research test tending to prefer problems not commonly preferred by the group. They further show that mediocrity of performance goes with mediocrity of preference as valued here, and that low performance corresponds with usualness of preference.

CONCLUSION

It may be said, then, that reported preference within a narrow field can be measured with a remarkable degree of accuracy by employing definite reporting methods in narrowed-down preference situations. Also, by comparing these measures with performance on the same items, a correlation to the degree of .56 was found. This coefficient of correlation, although significantly positive, does not warrant one glibly to predict the success of a pupil in solving arithmetic problems, the expressed preference of the pupil being known.

THE KOHS BLOCK-DESIGN TESTS. A REVISION FOR CLINICAL PRACTICE

MAX L. HUTT

Educational Clinic, College of the City of New York

PURPOSE OF THIS STUDY

In a recent article,[1] the writer introduced a technique for scoring the Kohs Block-Design Tests on the basis of time alone, thereby eliminating the more cumbersome, original method which involves the use of both time and moves as criteria. It was emphasized that the older scoring scheme had the following serious limitations: (1) It was highly subjective, leaving room for significant errors in scoring; (2) it precluded the observation by the examiner of the subject's behavior responses during the test; (3) the combined scoring of time and moves was not justified by the data presented by Kohs. The previous study was a preliminary report only, including but 33 examinees, and suggesting merely tentative conclusions. The revised method appeared to give at least as satisfactory results as the original method and at the same time obviated the above-mentioned difficulties.

Clinical psychologists are searching avidly for additional materials to supplement the existing tests in clinical use. If this revision of the Kohs test will render its administration more efficient and will supply a comparatively simple device for securing additional valuable data about the subject, it will have accomplished its primary purpose. The scope of the present research project is threefold. A corroboratory study will be made of the writer's revision of the scoring method. A critical evaluation will be made of the diagnostic values of the Kohs test in clinical practice for differentiating

[1] M. L. Hutt. "A Simplified Scoring Method for the Kohs-Block Designs Test." *Amer. J. Psychol.*, 42, 1930, 450–452.

various levels of mentality. Finally, a brief study will be made of the test's validity as a measure of general intelligence.

SELECTION OF SUBJECTS

The subjects were 100 boys and girls who had been examined at the Pediatric-Clinic of the Mt. Sinai Hospital in New York City.[2] The selection of cases was made by consulting the files of the Clinic and extracting the records of the first 100 individuals who had been given both the Kohs Block-Design Test and the Stanford-Binet Scale. Care was taken to eliminate all cases in which there was any doubt concerning the reliability of the results of either test. This precaution was necessary since the children referred to the Clinic presented problems in maladjustment. Chronologically they ranged from 6 years 4 months to 15 years 3 months. In intelligence as measured by the Stanford-Binet Scale they ranged from an IQ of 59 to an IQ of 155, the median IQ being 94.7.

THE REVISED SCORING TECHNIQUE

In his previous article[1] the writer explained the derivation of the revised scoring method. It is presented here in greater detail.

Table I presents both the old and the new techniques.

It will be noted that in the original scheme three factors, success, time or speed, and moves or accuracy, were rated. Consider two illustrations of this method. Design No. 1 has a score value of 3. This maximum score is attained if a subject completes the design in less than 21 seconds, and with less than 6 moves. If 21 or more seconds are required, one point is deducted from the score, and if 6 or more moves are made, an additional point is deducted. Similarly for design 4, if the subject requires 10 or more moves, one point is subtracted, for 31 seconds to 1 minute an additional point is sub-

[2] The writer is indebted to Dr. Ira S. Wile, director of the Clinic, who generously gave permission for the use of the records.

TABLE I

The original and revised scoring cards contrasted

DESIGN NO.	SCORE VALUE POINTS	POINTS TO BE SUBTRACTED		
		EXCESS TIME		EXCESS MOVES
		1 Point	2 Points	1 Point (Original Method)
		1.5 Points	3 Points	0 Points (Revised Method)
1	3	21″ and over		6 and over
2	5	31″ and over		7 and over
3	6	21″ to 35″	36″ and over	8 and over
4	7	31″ to 1′ 0″	1′ 1″ and over	10 and over
5	7	36″ to 1′ 5″	1′ 6″ and over	11 and over
6	7	36″ to 1′ 0″	1′ 1″ and over	12 and over
7	7	41″ to 1′ 10″	1′ 11″ and over	11 and over
8	8	41″ to 55″	56″ and over	10 and over
9	9	56″ to 1′ 10″	1′ 11″ and over	15 and over
10	9	1′ 56″ to 2′ 10″	2′ 11″ and over	22 and over
11	8	1′ 46″ to 2′ 30″	2′ 31″ and over	19 and over
12	9	2′ 21″ to 2′ 40″	2′ 41″ and over	30 and over
13	9	2′ 21″ to 2′ 33″	2′ 34″ and over	31 and over
14	9	2′ 26″ to 2′ 40″	3′ 41″ and over	32 and over
15	9	2′ 41″ to 3′ 0″	3′ 1″ and over	32 and over
16	10	2′ 41″ to 3′ 5″	3′ 6″ and over	31 and over
17	11	2′ 41″ to 2′ 55″	2′ 56″ and over	30 and over

tracted, and if 1 minute and 1 second or over are required, still another point is deducted. There is a maximum time limit for each design which is usually about 1 minute more than is necessary for its completion.[3] The typical division of points then is in the ratio 4:2:1 for success, speed, and accuracy respectively which Kohs proposed as the correct weighting of these elements.[4] This ratio was selected since, as Kohs demonstrated, the numerical values of the correlations of the three elements with life age were in the ratio 4:2:1. The writer contests the validity of this argu-

[3] For a more detailed consideration of the original scoring method consult, S. C. Kohs, *Intelligence Measurement*, 1923.

[4] Kohs, *op. cit.*, p. 43.

ment. Although the correlation coefficient between variable one and variable two is twice as great as the correlation coefficient between variable three and variable two (*i.e.*, success and life age compared with time and life age) the importance or value of variable one is not *necessarily* greater than or twice as great as the value of variable three. In addition, the writer has been unable to find in Kohs' data, any evidence that the inclusion of the scoring of moves has increased either the diagnosticity of the test or its validity.

The revised scoring method is now illustrated for the same designs, numbers 1 and 4. Design No. 1 retains its original score value of 3. The maximum score value is attained if the design is completed in less than 21 seconds, *regardless* of the number of moves. If 21 or more seconds are required, 1.5 points are subtracted, giving a score of 1.5. Similarly for design 4, which has a maximum score value of 7, if the subject requires 31 seconds to 1 minute, 1.5 points are deducted, and for 1 minute and 1 second or over, 3 points are deducted. Therefore, no consideration in scoring is given to the number of moves made in completing each design. It will now be purposeful to examine the results secured with both original and revised methods.

PROCEDURE AND RESULTS

The tests for the 100 cases had already been scored by the original method. The derived scores had been converted into Kohs Mental Ages (hereafter called Kohs M.A. 1) by the table of mental age equivalents given in the manual,[5] and the IQ's had been calculated. All of the tests were re-scored in accordance with the revised method; the newly derived scores were converted into Revised Kohs Mental Ages (here after called Kohs M.A. 2), by the table of mental age equivalents in the manual, and the IQ's were obtained.

Table II presents the Pearson correlation coefficients between the results of the original and the revised methods.

[5] Idem., p. 73.

TABLE II

Pearson correlation coefficients for 100 cases[6]

Stanford-Binet M.A. and Kohs M.A. (1)	$+.698 \pm .034$
Stanford-Binet M.A. and Kohs M.A. (2)	$+.708 \pm .033$
Stanford-Binet IQ and Kohs IQ (1)	$+.644 \pm .039$
Stanford-Binet IQ and Kohs IQ (2)	$+.630 \pm .040$
*Kohs M.A. (1) and Kohs M.A. (2)	$+.992 \pm .001$
*Kohs IQ (1) and Kohs IQ (2)	$+.994 \pm .001$

*In the writer's previous article (*op. cit.*, p. 451) a correlation of $+.93 \pm .01$ was secured between Kohs M.A. (1) and Kohs M.A. (2).

The obtained correlations between the Stanford-Binet M.A. and Kohs M.A. (1) and between the Stanford-Binet IQ and the Kohs IQ (1) may be compared with the results of other investigations. Kohs found a correlation for the mental ages of $+.82 \pm .01$ for 366 cases.[7] When these cases were divided into two groups of 291 public school cases and 75 feebleminded cases the correlations were $+.81 \pm .01$ and $+.67 \pm .05$ respectively. He obtained an r of $+.80$ between the Kohs and Binet IQ's for the entire group, an r of $+.58$ for the public school cases and an r of $+.67$ for the feebleminded. The discrepancy between the M.A. and IQ correlations is similar to that of this study. Broom reports an M.A. correlation of $+.65 \pm .08$ for 22 elementary school children.[8] Recently Wile and Davis made a thorough study of the Kohs test and obtained the following correlations for 250 cases: for IQ's, $+.63 \pm .02$; for M.A.'s, $+.82 \pm .01$.[9] Brown reported a correlation of $+.46 \pm .04$ between the Kohs M.A.'s and the National Intelligence Test M.A.'s.[10]

[6] The writer wishes to acknowledge the assistance of Samuel Goldberg in performing much of the statistical labors and in gathering most of the data.

[7] Kohs, *op. cit.*, pp. 181 ff.

[8] E. Broom. "The Validity of Four Individual Tests of Mental Ability." *Educ. Res. Bull.*, Los Angeles City Schools, 8, 1929, 9f.

[9] I. S. Wile and R. Davis. "A Comparative Study of the Kohs-Block Design Test." *Amer. J. Orthopsychiatry*, 1, 1930, 96.

[10] A. J. Brown. "An Enquiry into the Standardization of the Kohs Block-Design Test." JOURNAL APPLIED PSYCHOLOGY, 14, 1930, 178–181.

[In a study on a small group of foreign children Lester obtained a correlation of + .60 between Binet M.A. and Kohs M.A.][11] It is obvious that the size of the correlations between the Kohs M.A. and the Binet M.A. will vary somewhat with the group studied, but the evidence here presented is in general agreement with Kohs' original findings.

Further evidence concerning the comparability of the results of the old and the revised techniques is presented in tables III, IV, V, and VI.

TABLE III

Amount of change in months from Kohs M.A. (1) to Kohs M.A. (2)

Change	+5 to +6	+4	+3	+2	+1	0	−1	−2	−3	−4	−5
Frequency	3	2	22	13	10	27	6	4	7	3	3

TABLE IV

Amount of change from Kohs IQ (1) to Kohs IQ (2)

Change	+5	+4	+3	+2	+1	0	−1	−2	−3	−4	−5 to −6
Frequency	1	0	14	20	12	33	6	3	8	1	2

TABLE V

Percentile distribution of Kohs M.A. (1) and corresponding Kohs M.A. (2)

PERCENTILE	10	20	30	40	50	60	70	80	90
Kohs M.A. (1)	72	81	92	98	109	122	131	139	155
Kohs M.A. (2)	67	84	92	100	111	122	134	141	155

TABLE VI

Percentile distribution of Kohs IQ (1) and corresponding Kohs IQ (2)

PERCENTILE	10	20	30	40	50	60	70	80	90
Kohs M.A. (1)	58	66	77	81	85	92	100	111	121
Kohs M.A. (2)	58	68	79	82	85	92	101	113	121

[11] O. P. Lester. "Performance Tests and Foreign Children." *J. Educ. Psychol.*, 20, 1929, 304.

It is apparent that the middle 50 per cent of the distribution of changes for mental age range from 0 to about + 2 months, and for intelligence quotient ranges from 0 to + 2 points. The revised method of scoring, then, gives results which are highly comparable with the original method but tends to raise slightly both the mental age and the intelligence quotient. For clinical use the results of the two techniques may be regarded as equivalent. An additional check upon the equivalence of the two methods was made. The changes in months of mental age were examined for cases having the ten highest IQ's, the ten lowest IQ's, and ten average or normal IQ's. The arithmetic average of the changes for these three groups were respectively 2 months, 2 months, and 1.4 months, while the average change for the entire distribution of 100 cases was 1.9 months. It seems fair to conclude from the data presented that the revised scoring scheme is about equally efficacious for superior and inferior individuals. It is also efficacious for individuals with high mental ages and for those with low mental ages. One caution should be observed. It may be that with an extremely psychopathic patient the change in mental age and in intelligence quotient will be greater than any of the changes discovered in the present investigation. The response of such an individual will necessarily be regarded as pathologic, and his score, no matter how derived, will be considered invalid. Since the revised technique tends to eliminate the undue weighting of such extraneous elements as psychopathic instability, it may actually be the more valid.

In order to analyze the power of the Kohs test to differentiate the borderline group of individuals, those patients having IQ's between 70 and 79 inclusive were selected for separate study. The 22 cases thus selected showed a correlation of $+ .13 \pm .15$ between Binet IQ and Kohs IQ (2). Superficially, it appears that the Kohs test fails to differentiate the borderline group. This, indeed, is the conclusion of Wile and Davis who reported a correlation of + .02 between

Kohs IQ and Binet IQ for 45 borderline cases.[12] The unreliability of the correlation coefficients, as judged by their probable errors, and the lack of other adequate criteria, since the Binet test is not an infallible measure of intelligence especially for maladjusted children, seem to the writer sufficient grounds for vitiating the "apparent" conclusion. A more detailed account of the Kohs IQ's for the 22 borderline cases of the present study is therefore displayed in table VII.

TABLE VII

Distribution of Kohs IQ's for 22 cases having Binet IQ's between 70 and 79

Kohs IQ	45–49	50–54	55–59	60–69	70–79	80–89	92	105
No. of cases	2	0	6	5	3	3	2	1

It will be noted that only three individuals obtained IQ's higher than 89, and that in general for the entire borderline group the effect of the Kohs test is to depress the IQ rating. Since the "dysgenic" environmental factors, such as language difficulty, home environment, and the like, appear to have less effect on the Kohs test than on the Binet, the former may prove even more helpful than the latter in differentiating the borderline cases among maladjusted individuals. The present data do not indicate whether either test alone, or both tests used in conjunction, can differentiate adequately children of the borderline group.

It is useful to judge the validity of the Kohs test by means of the "demarcation of age groups" principle. Brown has already given this point some consideration.[13] It may be profitable to compare his findings with the results of this study. Ages 9, 10 and 11 are used since these ages have the greatest frequencies. Brown examined the median raw score at each age, and the arithmetical deviations of the scores at each age level. This writer has used the median mental age

[12] Wile and Davis, *op. cit.*, p. 97.

[13] Brown, *op. cit.*, p. 180.

in months, and the quartile deviation expressed in months. Similar data are provided for the Binet test. Table VIII contains the essential figures.

TABLE VIII

Age demarcations of the Kohs and Binet tests compared

Chronological Age	9 (15)*	10 (13)	11 (16)
Median C.A.	9–5	10–7	11–5
Kohs M.A.	114 ± 9.0	110 ± 13.0	142 ± 28.0
Binet M.A.	122 ± 18.0	100 ± 26.5	122.5 ± 35.0
Brown's Data	16 ± 13.0	20 ± 15.0	27 ± 18.0

* The figures in parentheses represent frequency.

The data indicate marked overlapping both on the Kohs test and on the Binet test. The writer used the quartile deviation rather than the arithmetical deviations employed by Brown, in view of the greater reliability of the former for the present data. The small number of cases at each age level limits the conclusiveness of the possible interpretations. Brown concluded,[14] "Judged by the principle of demarcation of age groups . . . the validity of the Kohs test is of doubtful clinical value." If this were so, then by the same line of reasoning, the data presented in table VII would render the Binet "of doubtful clinical value." The validity of the Binet test has long been assumed. If reliance is placed in that assumption the Kohs test may still show sufficient validity when judged by the age demarcation principle since for the present data at least it shows less overlapping than the Binet. The increase in median mental age at each age level and the relatively small quartile deviations, despite large variations in intelligence quotient, tend to confirm rather than negate the thesis that the test measures well whatever it does measure.

[14] Idem., p. 181.

SUMMARY AND CONCLUSIONS

1. The present study offers highly confirmatory evidence that the revised scoring method may be substituted for the original method. The scoring of moves is eliminated, and the administration of the test becomes more efficient and time-saving.

2. The revised technique is about equally reliable for both high and low mental ages, and for both high and low intelligence quotients.

3. The most probable changes in mental age tend to fall between 0 and + 2 months. The average change in mental age is + 1.9 months.

4. The most probable changes in intelligence quotient tend to fall between 0 and + 2 points.

5. Caution should be observed in the interpretation of the test results of psychopathic individuals.

6. There is no palpable evidence that scoring the Kohs test on the basis of time alone lowers the test's diagnostic value.

7. The test appears to measure well whatever it does measure; but large discrepancies may be expected between the Kohs test and the Binet test results in some cases.

8. The Kohs test may prove especially helpful in diagnosing many borderline cases, since it does not appear to be markedly affected by dysgenic environmental factors.

9. An additional study should be made of those cases showing large discrepancies between Binet and Kohs scores.

A CONSIDERATION OF THE DIFFERENCES IN READINESS OF RECALL OF THE SAME ADVERTISEMENTS BY MEN AND WOMEN[1]

KEY LEE BARKLEY
University of Illinois

The following study was made incidentally in the course of a larger research. It was discovered that the group of subjects, whose results are reported in this paper, was distributed as indicated below in age, sex, and in classification in the university. (There were some few additional cases the results from which were not included.) This discovery led to a consideration of the problem indicated in the title of the article. The results of the more extensive research are to appear in other writings.

PURPOSE

To discover whatever differences there might be in the readiness of men and women later to recall identical advertisements which had been presented to them at the same time and place.

SUBJECTS

The group of subjects was composed of forty-two university students, twenty-one men and twenty-one women. They were paired in age and in classification in the university; that is, for each man there was a woman of the same age and classification, except in three cases. In these three cases the difference was only in age and was not more than two years in any case. There were twelve first year students, fourteen second

[1] Reported from the University of North Carolina Psychology Laboratory. This research was done under the direction of Dr. Harry W. Crane.

year, fourteen third year, and two fourth year students. Their ages ranged from seventeen to twenty-two years.

MATERIALS

An issue of the *Saturday Evening Post* containing one hundred twenty-three advertisements was used for study. The magazines were secured for use before the subjects had an opportunity to see them.

PROCEDURE

The *Saturday Evening Posts* were brought into the class room at the time the students were entering for the regular class scheduled for that hour. The regular instructor informed the students that it had been arranged to have them cooperate in a bit of psychological research and asked them to give their intelligent cooperation.

The magazines were distributed to the subjects who were asked to let them lie on the arm of the chair face up until they were given instructions as to the manner of procedure. The experimenter then gave the following instructions:

Place your *Saturday Evening Post* before you face up.

When I say "turn," you are to turn back the front cover. You are to spend the time until I next say "'turn" in looking at the two exposed pages in any way that interests you—trying, however, to do it in as natural a manner as you would if you were seated in your own room.

Please turn to the next page promptly each time the signal "turn" is given.

Pay no attention to your neighbors.

The time of exposure of each two pages and also for the back cover was fifteen seconds.

When the subjects had turned all the pages, the magazines were taken up and placed where the subjects could not see them. University quiz books were then given the subjects with the following instructions:

Write down as quickly as possible all the advertisements that you can recall, being sure to write them down just as you happen to remember them. About each advertisement recalled try to give as accurately as possible the following two things in the order named:

1. The definite name of the article advertised.
2. The one thing in the advertisement that most stands out for you.

Twenty-two minutes were allowed for recalling the advertisements.

RESULTS AND DISCUSSION

In compiling the data the recalls listed in the quiz books were identified and given the proper number from the magazine. In order to get the total number of recalls and the sum of the ranks of recalls for each advertisement, all the identified advertisements were distributed on large sheets of ruled paper. The number of the advertisement was indicated in a vertical column at the left and the rank of recall in a horizontal column at the top of the sheet.

The indices of efficiency of the various advertisements were then found by employing the following formula:[2]

$$\frac{N^2}{\text{Sum R}} \times \frac{100}{gn} = E$$

N = Number of recalls.
R = Rank of recalls.
gn = Group number (number of subjects).
E = Efficiency.

The manner in which the formula is applied is illustrated below:

To get the sum of the ranks multiply the rank of recall by the number of recalls under that rank and take the sum of the products for all ranks. Thus, suppose we have a group of one hundred subjects and that a certain advertisement is listed in

[2] A criticism of this formula may be found in the Journal of Applied Psychology, Vol. XVI, No. 1, Pages 74–90.

places 3, 6, and 9 by several subjects. Let us consider that under rank three there are three recalls, under rank six two recalls, and under rank nine one recall. The sum of the ranks would then be

$$(3 \times 3) + (2 \times 6) + (1 \times 9) = 30$$

The number of recalls would be $3 + 2 + 1 = 6$.

Since there are 100 subjects "gn" would be 100.

We could then substitute and solve the formula as follows:

$$E = \frac{(6)^2}{30} \times \frac{100}{100} = \frac{36}{30} = 1.2000.$$

The first step in treating the results obtained from the two equilibrated groups was to find the correlation between the indices of efficiency of all the advertisements as determined for the men and for the women. The Pearson Product Moment method was used. A correlation of plus $.737 \pm .0245$ was found. This correlation is regarded as indicating that there was no significant difference between the readiness of the men and women to react to most of the advertisements in the group. This result would tend to substantiate the conclusion stated by Adams[3] that there are no significant differences between the reactions of men and women to advertisements of commodities of general use and that results obtained from men and women subjects may be treated together without serious error.

To discover if there were classes of commodities advertisements of which were recalled more readily by either men or women, two tables were made, one showing the twenty advertisements with the highest indices of efficiency for men, and the other the twenty advertisements with the highest indices for women. The following summary table was made from the itemized tables to show the number of times certain commodities were mentioned.

In order to indicate the general agreement between men and women in readiness to recall these selected advertisements, the

[3] Adams, H. F., Adequacy of Laboratory Tests in Advertising. Psychological Review, 22, 1915, 402–422.

COMMODITY	NO. OF TIMES MENTIONED BY WOMEN	NO. OF TIMES MENTIONED BY MEN
1. Automobiles and accessories	9	11
2. Food	3	2
3. Clothing	2	4
4. Personal	2	1
5. Household	3	0
6. Movies	1	0
7. Business or industry	0	2

correlation between the indices of efficiency of all the advertisements which appeared in the two tables of twenty was worked out. There was a total of twenty-eight different advertisements. The Pearson Product Moment method was again used. The coefficient of correlation in this case was only plus .325 with a probable error of $\pm$ .114. This result indicates that some factor or factors operated more strongly in the case of these selected advertisements than in the group as a whole to make a difference in the readiness of men and women to recall advertisements of certain classes of commodities.

Since these classes of commodities were not clearly defined by selecting the twenty advertisements with the highest indices of efficiency for men and for women, two other tables were made which present the advertisements for which the indices of efficiency for the women were at least one efficiency unit greater than the indices for men, or vice versa. That is, if one advertisement had an index of 4.5000 for women, then it would have to have an index of 5.5000 or greater to be selected for the table showing the men's preferences. Each table happened to contain twenty-two advertisements. From these itemized tables a summary table (see p. 313) was made to show the number of times advertisements of the given commodities appeared in the lists.

The results given in the summary table indicate that there are quite marked sex differences with respect to advertisements of certain types of commodities. Men more frequently and readily recall advertisements concerning automobile accesso-

COMMODITY	MEN	WOMEN
1. Automobiles	2	3
2. Automobile accessories—spark plugs, tires, etc.	6	1
3. Clothing for men	5	1
4. Clothing for all the family	0	3
5. Business or industry—paint, tools, motors, etc.	6	2
6. Food	0	3
7. Household, rugs, furnishings, furniture polish, etc.	0	7
8. Moving pictures	0	2
9. Shaving	3	0

ries, clothing for men, business or industry, and materials for shaving. Women more readily and frequently recall advertisements concerning food, household furnishings, moving pictures, and clothing for all the family.

The differences in the readiness and frequency of recall of the advertisements of these types of commodities by men and women may be explained by a difference in the life history of men and women. There is nothing to suggest that the different advertisements take advantage of a difference in native equipment; the same general form, the same sizes, and the same method of presentation are used for both men and women. The difference must lie in the acquired preparation of the individuals to react to the advertisements of certain commodities. By different training and exercise men and women are differently modified so that a given stimulus may be more potent for one than for the other. Men and women do not come into contact with nor use the different classes of commodities to the same extent, hence a difference in modified structure probably is present. This difference in contact and use should be especially apparent upon considering the training of the subjects used in the experiment with regard to men's clothing, auto accessories, building materials, commodities of special significance to the housekeeper, and materials for shaving.

There is no distinct difference between men and women with respect to advertisements of automobiles.

The number of subjects used in this experiment, and the range of types of commodities represented in the list of advertisements were not great enough to justify a final statement on this matter of sex differences, but they are suggestive of the relation which probably does exist.

CONCLUSIONS

1. There is no significant difference in the readiness of men and women to recall advertisements of commodities of general or common use, provided they have observed them under like conditions.

2. The results obtained from men and women subjects may be treated together unless the advertisements are of selected commodities.

3. There is a notable sex difference in readiness to recall advertisements of certain selected commodities. This difference depends upon a difference in the training and activities of men and women.

BIBLIOGRAPHY

1. Adams, H. F., Adequacy of Laboratory Tests in Advertising. Psychol. Rev., 22, 1915, 402–422.
2. Adams, H. F., Advertising and Its Mental Laws. The Macmillan Co., New York, 1916.
3. Franken, Richard B., Advertising Appeals Selected by the Method of Direct Impression. J. Appl. Psychol., 8, 1924, 232–244.
4. Holmes, Joseph L., The Free Association Method as a Measure of the Efficiency of Advertising. J. Appl. Psychol., 9, 1925, 60–65.
5. Hollingworth, H. L., Judgments of Persuasiveness (of Advertising Appeals). Psychological Review, 18, 1911, 234–256.
6. Naether, Carl A., Advertising to Women. Prentice-Hall, New York, 1928.
7. Poffenberger, A. T., The Return Coupon as a Measure of Advertising Efficiency. J. Appl. Psychol., 7, 1923, 202–208.
8. Poffenberger, A. T., Psychology in Advertising. A. W. Shaw Co., New York, 1928.
9. Poffenberger, A. T., and Franken, R. B., A Study of the Appropriateness of Type Faces. J. Appl. Psychol., 7, 1923, 312–329.
10. Starch, Daniel, Principles of Advertising. A. W. Shaw Co., New York, 1926.

COMPETITION FOR MARKS

JOHN W. CHARLES
Iowa State Teachers College

A survey of the marks of one hundred 1930 spring term classes in Iowa State Teachers College was made by the writer to ascertain the degree to which marks were distributed on the basis of the classes to which the students belonged. The one hundred classes represented sixty-five different courses given in twelve different departments of college work. No class was selected which did not contain students of at least two different collegiate years and none was selected which did not contain at least fourteen students. The following tables show the combinations of classes, the frequency of different sized classes, the departments represented in the different classes, and the distribution of students by college years.

TABLE 1

Freshmen and Sophomores	31 classes
Freshmen, Sophomores, and Juniors	29 classes
Sophomores, Juniors, and Seniors	19 classes
Juniors and Seniors	13 classes
Freshmen, Sophomores, Juniors, and Seniors	5 classes
Freshmen, Sophomores, and Seniors	2 classes
Freshmen, Juniors, and Seniors	1 class
Total	100 classes

The marking system of Iowa State Teachers College is based on a five level distribution using the symbols A, B, C, D, F, in which F represents failure. A suggestion is made to the faculty that in large classes at least, the proportion of marks should approximate A, 7 per cent; B, 24 per cent; C, 38 per cent; D, 24 per cent; F, 7 per cent. It is a matter of interest and some concern to know just which students are

TABLE 2

10 to 14 enrollment	1 class
15 to 19 enrollment	15 classes
20 to 24 enrollment	30 classes
25 to 29 enrollment	19 classes
30 to 34 enrollment	6 classes
35 to 39 enrollment	13 classes
40 to 44 enrollment	11 classes
45 to 49 enrollment	4 classes
50 to 54 enrollment	0 classes
55 to 59 enrollment	1 class
Total	100 classes

TABLE 3

Natural Sciences	14 classes
Music	5-classes
Physical Sciences	3 classes
Psychology and Education	20 classes
English	20 classes
Mathematics and Commerce	5 classes
Modern Languages	4 classes
Art and Drawing	3 classes
Religious Education	5 classes
Physical Education	2 classes
Teaching	3 classes
Social Sciences	16 classes
Total	100 classes

TABLE 4

Freshmen Students	985
Sophomore Students	987
Junior Students	390
Senior Students	479
Total	2841

obtaining the superior marks and which are obtaining the inferior marks when students of different collegiate ranking are enrolled in the same courses. The following tables show the

total number of marks secured by students of the four college classes and the per cent of marks secured by each of the classes.

Number of times each mark was obtained by Freshmen and Sophomores in the same courses. Thirty-one classes

	A	B	C	D	F	TOTAL
Freshmen	28	111	235	92	17	483
Sophomores	13	100	157	34	8	312
Total	41	211	392	126	25	795

Per cent of marks obtained by each class in the same thirty-one classes

	A	B	C	D	F	PER CENT OF THE GROUP
Freshmen	68.3	52.6	59.9	73	68	60.8
Sophomores	31.7	47.4	40.0	27	32	39.2
Total	100	100	99.9	100	100	100

This is to be read as follows. The Freshmen, though they were in numbers 60.8 per cent of the 31 classes, obtained 68.8 per cent of the A's, etc.

Number of times each mark was obtained by Freshmen, Sophomores, and Juniors in the same courses. Twenty-nine classes

	A	B	C	D	F	TOTAL
Freshmen	20	95	172	82	25	394
Sophomores	15	110	186	48	13	372
Juniors	8	27	41	7	2	85
Total	43	232	399	137	40	851

Per cent of marks obtained by each class

	A	B	C	D	F	PER CENT OF THE GROUP
Freshmen	46.5	40.9	43.1	59.9	62.5	46.3
Sophomores	34.9	47.4	46.6	35.0	32.5	43.7
Juniors	18.6	11.6	10.3	5.1	5.0	10.0
Total	100	99.9	100	100	100	100

Number of times each mark was obtained by Sophomores, Juniors, and Seniors in the same courses. Nineteen classes

	A	B	C	D	F	TOTAL
Sophomores	9	39	123	28	8	207
Juniors	12	37	62	25	5	141
Seniors	15	83	101	19	0	218
Total	36	159	286	72	13	566

Per cent of marks obtained by each class

	A	B	C	D	F	PER CENT OF THE GROUP
Sophomores	25	24.5	43	38.9	61.5	36.6
Juniors	33.3	23.3	21.7	34.7	38.5	24.9
Seniors	41.7	52.2	35.3	26.4	0.0	38.5
Total	100	100	100	100	100	100

Number of times each mark was obtained by Juniors and Seniors in the same courses. Thirteen classes

	A	B	C	D	F	TOTAL
Juniors	11	36	79	17	3	146
Seniors	16	74	107	40	5	242
Total	27	110	186	57	8	388

Per cent of marks obtained by each class

	A	B	C	D	F	PER CENT OF THE GROUP
Juniors	41.0	32.7	42.5	29.8	37.5	37.6
Seniors	59.0	67.3	57.5	70.2	62.5	62.4
Total	100	100	100	100	100	100

Number of times each mark was obtained by Freshmen, Sophomores, Juniors, and Seniors in the same courses. Five classes

	A	B	C	D	F	TOTAL
Freshmen	1	2	26	16	6	51
Sophomores	3	23	36	12	4	78
Juniors	2	6	8	0	0	16
Seniors	1	6	5	0	0	12
Total	7	37	75	28	10	157

Per cent of marks obtained by each class

	A	B	C	D	F	PER CENT OF THE GROUP
Freshmen	14.3	5.4	34.7	57.1	60	32.5
Sophomores	42.9	62.2	48.0	42.8	40	49.7
Juniors	28.6	16.2	10.7	0.0	0	10.2
Seniors	14.3	16.2	6.7	0.0	0	7.6
Total	100.1	100.0	100.1	99.9	100	100.0

Number of times each mark was obtained by Freshmen, Sophomores, and Seniors in the same courses. Two classes

	A	B	C	D	F	TOTAL
Freshmen	1	9	12	18	3	43
Sophomores	0	9	7	2	0	18
Seniors	0	2	0	0	0	2
Total	1	20	19	20	3	63

Per cent of marks obtained by each class

	A	B	C	D	F	PER CENT OF THE GROUP
Freshmen	100	45	63.2	90	100	68.3
Sophomores	0	45	36.8	10	0	28.6
Seniors	0	10	0.0	0	0	3.2
Total	100	100	100.0	100	100	100.1

Number of times each mark was obtained by Freshmen, Juniors, and Seniors in the same course. One class

	A	B	C	D	F	TOTAL
Freshmen	1	4	7	2	0	14
Juniors	1	1	0	0	0	2
Seniors	3	1	1	0	0	5
Total	5	6	8	2	0	21

Per cent of marks obtained by each class

	A	B	C	D	F	PER CENT OF THE GROUP
Freshmen	20	66.7	87.5	100	0	66.7
Juniors	20	16.7	0.0	0	0	9.5
Seniors	60	16.7	12.5	0	0	23.8
Total	100	100.1	100	100	0	100

Let the marks A and B be called superior marks and the marks D and F be called inferior marks. Therefore there are two chances in each course considered for a class to obtain more or less than its proportion of superior or inferior marks. The tables above will show that the Freshmen in competition with other classes obtained more than their proportion of superior marks sixty-two times out of a possible one hundred thirty-six, and that they received more than their proportion of inferior marks one hundred thirty-five times out of a possible one hundred thirty-six. In competition with Juniors and Seniors the Freshmen received all of the inferior marks; in competition with Sophomores and Seniors the Freshmen received all of the F's and nine-tenths of the D's; in competition with Sophomores, Juniors, and Seniors the Freshmen received a trifle over one-half of the inferior marks; in competition with Sophomores and Juniors the Freshmen received three-fifths of the inferior marks; in competition with the Sophomores the Freshmen received just short of two-thirds of the inferior marks.

Seniors, in contrast, received more than their proportional share of superior marks sixty-four times out of a possible eighty times, and more than their proportional share of inferior marks twenty-six times out of a possible eighty. Both Sophomores and Juniors secured more than their share of superior marks and less than their share of inferior marks throughout the study. The Sophomores made their best record at the B level. A complete tabulation follows (see p. 321) showing exactly how often each class received more than its proportional share of superior and inferior marks.

Previous studies made at Iowa State Teachers College by Dr. M. J. Nelson have revealed the difficulties encountered by those students who receive inferior marks. In the fall term 1929 three-fourths of the students failing one or more courses fell in the lower half of the scores made in the Freshman testing program, and just about half of the students who made failing marks fell in the lowest twenty-five per cent of marks

Freshmen:									
	Superior	marks,	62	times	out	of	a	possible	136 times.
	Inferior	marks,	135	"	"	"	"	"	136 "
Sophomores:									
	Superior	marks,	117	"	"	"	"	"	172 "
	Inferior	marks,	38	"	"	"	"	"	172 "
Juniors:									
	Superior	marks,	102	"	"	"	"	"	134 "
	Inferior	marks,	38	"	"	"	"	"	134 "
Seniors:									
	Superior	marks,	64	"	"	"	"	"	80 "
	Inferior	marks,	26	"	"	"	"	"	80 "

in the testing program. In the same study it was found that of the total number of Freshmen in the lowest ten per cent in the testing program practically thirty per cent withdrew from school, and practically seventy-five per cent either withdrew from school or failed to meet the standard for graduation. Students of low native ability certainly find it difficult to compete with students of higher classification than their own. Students of high native ability seem to get along in any classification. In view of these facts there is evident justification for the practice of closing some courses against Freshman participation.

ON THE RELIABILITY AND VALIDITY OF DERIVED SCORES YIELDED BY THE McCALL MULTI-MENTAL SCALE

J. C. LAUDERBACH AND ENID HAUSE
Public Schools, Chula Vista, California

It is the purpose of this report to present data concerning the reliability (constancy) of the IQ yielded by the McCall Multi-Mental Scale and the validity of the mental age score yielded by this test. The measures which form the subject matter of this report were obtained during October, 1929, and September, 1930, in the Chula Vista, California, Elementary School.

One hundred fifty pupils in grades 4–6 took the McCall Multi-Mental Scale twice. The Pearson r representing the degree of relationship between the two paired series of scores is positive .791; P. E.$_r$, $\pm$ .021. The two series of measures may be summarized as follows: The range in IQ score values on the first (October, 1929) testing was from 64 to 162 with a mean score of 107.4 and a standard deviation of 18.4; on the second (September, 1930) testing, from 62 to 168, with a mean score of 105.8 and a standard deviation of 20.3.

Fifty-seven pupils had IQ scores higher on the second test than on the first. The score range of these increased scores extended from 1 to 45 points IQ, with a median of 9.7 points IQ higher on the second test than on the first, and a quartile deviation of 5.7 points IQ. Ninety pupils had IQ scores lower on the second test than on the first. The score range of these decreased scores varied from 1 to 26 points IQ, with a median of 7.6 points IQ lower on the second test than on the first, and a quartile deviation of 3.5 points IQ. Only three pupils in the group of 150 made exactly the same IQ score on both the first and second tests. The findings reported here do not

agree with Jordan's statement:[1] "There is little change in the average IQ. . . . The trend seems to vary upward slightly although there are one or two exceptions to this rule." In the present investigation it has been shown that the trend in IQ shift is downward on the re-test.

Three hundred twenty-four pupils in grades 4–6 took both the New Stanford Achievement Test, Advanced Examination, and the McCall Multi-Mental Scale during the school years of 1929–1930 and 1930–1931. The Pearson r representing a measure of mutual validity between the paired mental and educational age scores is positive .752; P. E., $\pm$ .016. The two series of measures may be summarized as follows: the New Stanford educational age score ranged between 7–5 and 15–2, with a mean of 10–2 and a standard deviation of 1–4; the McCall Multi-Mental mental ages ranged between 6–6 and 19–2, with a mean of 10–9 and a standard deviation of 1–9.[2] It would seem that there is little need for using such a mental test as the McCall Multi-Mental Scale in grades 4–6, since apparently the test contributes little that we could not discover through using the achievement battery findings alone.

These findings are reported in the hope that they will elicit research by others in the same direction.

[1] Jordan, A. M. "Mental Growth." JOURNAL OF APPLIED PSYCHOLOGY, 14: 517–31. Dec., 1930.

[2] This finding agrees closely with one obtained (but not yet published) by M. E. Broom of the University of Southern California. He finds paired average grade placement scores yielded by the public school achievement test and McCall mental age scores to give a Pearson r of positive .805, 575 cases. His data were obtained in another small city school system in Southern California.

A ROUTINE PROCEDURE FOR OBTAINING COMPARABLE SCORES

PAUL HORST

U. S. Civil Service Commission, Research Division

The importance of reducing raw scores from two or more tests to some sort of comparable measures is quite generally recognized. The method most commonly used is that of converting raw scores to standard scores by the formula

$$Z = \frac{X - M}{\sigma}$$

where

Z = standard score.

X = raw score.

M = mean of the distribution of raw scores.

σ = standard deviation of the distribution of raw scores.

In other words the distributions of raw scores are converted to new distributions in which the means are equal to zero and the standard deviations are equal to unity.

Another method for making scores from several different tests comparable is to equate percentiles. That is, scores are regarded as equivalent if their percentile values are equal. Statistically this method is more general than the method of equating standard scores. If the distributions of raw scores are geometrically similar this method gives the same results as equating standard scores. In this case the relationship between percentile scores will be linear. However, if the distributions are geometrically dissimilar the relationship between percentile scores will be nonlinear. In general, of course, the relationships are not smooth, whether linear or nonlinear, due to the irregularity of the original distributions.

It may be shown mathematically that if the product moment correlation coefficient between two variables whose distributions are geometrically dissimilar has an absolute value

greater than zero that the absolute value of this correlation will be higher if the distributions are converted to geometrically similar forms. The reason for this is obvious. If the two distributions are geometrically similar then the relationship between the two variables will be linear. If the two distributions are geometrically dissimilar then there will be a curvilinear relationship between the two variables. The ordinary product moment correlation coefficient assumes a linear relationship between the correlated variables. Hence a correlation coefficient calculated on the assumption of a linear regression when the regression is actually curvilinear will be smaller in absolute value than if it were calculated on the assumption of curvilinearity.

Logically, then, it would seem advisable to reduce all distributions to the same form so that the interrelationship of the scores for all the distributions would be linear. This amounts to equating percentiles of the original distributions to percentiles of the common distribution. That is, the distribution selected as the one to which all others are to be reduced.

Perhaps the most obvious form of distribution to select for the one to which all the distributions are converted is the well known normal distribution.

From tables of the normal probability integral we can readily find base line values for any given percentile, that is, for any given proportion of area. These base line values will be given in the form of standard measures. Thus our problem amounts to equating percentiles from raw distributions to percentiles of a normal distribution with a mean of zero and a standard deviation of unity.

Aside from the fact that normal probability tables are generally available, the adoption of the normal probability curve as the form of the common distribution presents other less obvious, though vastly more significant advantages. Mathematically, the problems of multiple correlation are much more easily handled when all of the variables are normally dis-

tributed than when some or all of the distributions deviate from the normal curve. If all the variables considered are normally distributed the multiple correlation surface can always be represented by a definite mathematical function. Thus far it has not been possible to develop a generalized mathematical expression for correlation surfaces which are not normal. Even if such a form were developed it would be extremely complicated and quite impractical for application to experimental data.

Now, in order to equate percentile values of the distributions of raw scores to percentile values of a normal distribution we attempt to determine a function of the raw scores which will convert them into normally distributed scores. We select a simple power series as the function. Thus

$$Z = C_0 + C_1X + C_2X + C_3X + C_4X,$$

where

$Z =$ the transformed score.

$X =$ the raw score.

C_0, C_1, C_2, C_3, C_4 are constants to be determined.

We assume that a fourth degree equation gives us a sufficiently close approximation. There are five unknowns or constants in the transformation equation which must be determined. For the solution of any particular problem it is not necessary to understand the theory underlying the method. The theory is outlined in another article.[1] It is necessary only to set up the work sheet as indicated.

In column 1′ section A write the class interval limits of the distribution. There must be exactly eight class intervals. As a rule the range will not be exactly divisible by eight. Suppose the range is L. Then let us say

$$\frac{L}{8} = s + \frac{d}{8}$$

It is always possible to choose s so that the absolute value of d is not greater than 4. If d is even, $\frac{d}{2}$ is added to the first and

[1] Obtaining Comparable Scores from Distributions of Dissimilar Shape. Journal of the American Statistical Association, Dec. 1931. p. 455.

last class intervals of the distribution. If d is odd, $\frac{d-1}{2}$ and $\frac{d+1}{2}$ are added to the extreme intervals respectively, so that the interval having the lowest frequency will be the shorter of the two.

In column 2′ write the frequencies for each class interval. The frequencies which we use for illustration were taken from a distribution of intelligence test scores. In column 3′ cumulate the frequencies given in column 2′. In column 4′ convert these cumulative frequencies into percentage values. In column 5′ write the sigma values of these percentages as determined from a table of the normal probability integral.

Columns 1 to 5 inclusive of section A are the values which have been determined by the writer and remain the same for any problem. They may be copied directly from the above table.

Column S of section A gives horizontal summations of columns 1 to 5. This column also remains unchanged and may also be copied directly.

Columns 1 to S inclusive through lines 1 to 7 inclusive of section B are obtained by multiplying columns 1 to S inclusive of section A by column 5′. The values in these columns will depend on the values in column 5′ which in turn are derived from the particular distribution which is being transformed.

The values in line 8 of section B are constants which are copied directly from the table.

Line 9 of section B gives the vertical summations of column 1 to S in section B. If the multiplications and additions thus far have been correct the algebraic sum of the first 5 entries in line 9 of Section B should equal the last entry to within several points in the last decimal.

In section C the diagonal of 1's is constant for any problem. The signs in this section are also the same for any problem. The other values in section C are determined as follows:

We let A be the midpoint of the range and B the length of the class interval. Then in line 1 column 2 write $\frac{A}{B}$, that is, the mid-point of the range divided by the length of the class interval. In column 3, square this value. In column 4 cube it. In column 5 raise it to the 4th power. Symbolically the first line will be

$$1, -\left(\frac{A}{B}\right), +\left(\frac{A}{B}\right)^2, -\left(\frac{A}{B}\right)^3, +\left(\frac{A}{B}\right)^4$$

the second

$$1, -2\left(\frac{A}{B}\right), +3\left(\frac{A}{B}\right)^2, -4\left(\frac{A}{B}\right)^3$$

the third

$$1, -3\left(\frac{A}{B}\right), +6\left(\frac{A}{B}\right)^2$$

the fourth

$$1, -4\left(\frac{A}{B}\right)$$

the fifth

$$1$$

The values in the rows subsequent to row 1 may be readily calculated by putting first $\frac{A}{B}$ in the machine and multiplying sucessively by 2, 3, and 4, making the entries diagonally downward toward the right. Then $\left(\frac{A}{B}\right)^2$ is put in the machine and multiplied successively by 3 and 6. Finally $\left(\frac{A}{B}\right)^3$ is multiplied by 4 and entered in line 2 column 5, completing the section down to line 6. Line six consists of the values

$$\left(1-\frac{A}{B}\right)^5, \left(1-\frac{A}{B}\right)^4, \left(1-\frac{A}{B}\right)^3, \left(1-\frac{A}{B}\right)^2, \left(1-\frac{A}{B}\right)$$

The successive vertical summations of the columns in section C should equal the values in line 6 respectively to within several points in the last decimal.

Next multiply each column in section C by the value immediately above it in line 9 of section B and enter the products in column 1 to 5 inclusive of section D as indicated.

Column S of section D is a horizontal summation of the preceding columns in section D. The vertical summation of the first 5 entries in column S, section D, should equal the horizontal summation of the first 5 entries in line 6 of section D to within several points in the last decimal.

In column B, section D, make the successive entries

$$\left(\frac{1}{B}\right)^0, \left(\frac{1}{B}\right)^1, \left(\frac{1}{B}\right)^2, \left(\frac{1}{B}\right)^3, \left(\frac{1}{B}\right)^4$$

where B is the length of the class interval. Then multiply column S by column B and write the products in column Ci of section D. These values give the coefficients in the transformation equation. They are the constants required to transform the values in the original distribution into values which yield a normal distribution with a mean of zero and a standard deviation of unity. Thus, for the distribution given by column 2′ of section A the transformation equation is

$$Z = -4.0515 + .1963\,X - .00282\,X^2 + .0000148\,X^3 + .000{,}000{,}0839\,X^4.$$

To transform an entire distribution of values, a transmutation table may be readily constructed. The range of the distribution of original raw scores is known. The first four columns of the work sheet will contain the first four powers of the series. These may be copied from any table of powers of numbers. The next four columns are simply the first four columns multiplied by C_1, C_2, C_3, C_4, respectively. The ninth column is the horizontal summation of the second set of four columns plus the constant C_0. Thus the first column will give the original raw score values while the ninth column will give the desired transformed values.

It should be noted that the method above outlined is applicable only to comparatively large bodies of data. It should not be employed with less than three or four hundred cases, unless it is quite evident that the shape of the distribution would not be materially changed by increasing the size of the sample.

WORK SHEET FOR CALCULATING THE CONSTANTS OF THE TRANSFORMATION EQUATION

		1′	2′	3′	4′	5′	1	2	3	4	5	S	B	C_i
	1	6–13	24	24	.041	–1.73	–.128208	– .119526	.108102	.005892	–.006119	–.139859		
	2	14–21	72	96	.166	– .97	.069934	– .162461	.043996	.010943	–.003205	–.040793		
	3	22–29	134	230	.397	– .26	.314673	– .106063	– .061474	.007576	.002622	.157334		
A	4	30–37	128	358	.618	.30	.417256		– .107813		.005245	.314688		
	5	38–45	114	472	.815	.90	.314673	.106063	– .061474	–.007576	.002622	.354308		
	6	46–53	60	532	.919	1.40	.069934	.162461	.043996	–.010943	–.003205	.262243		
	7	54–61	33	565	.976	1.98	–.128208	.119526	.108102	–.005892	–.006119	.087409		
	8	62–69	14	579										
	1						.221800	.206780	– .187016	–.010193	.010586	.241956		
	2						–.067836	.157587	– .042676	–.010615	.003109	.039569		
	3						–.081815	.027576	.015983	–.001970	–.000682	–.040907		
B	4						.125177		– .032344		.001574	.094406		
	5						.283206	.095457	– .055327	–.006818	.002360	.318877		
	6						.097908	.227445	.061594	–.015320	–.004487	.367140		
	7						–.253852	.236661	.214042	–.011666	–.012116	.173070		
	8							– .434342		.070707		–.363635		
	9						.324588	.517164	– .025744	.014125	.000344	.830476		
	1						1.	–4.750000	22.5625	–107.1719	509.0664			
	2							1.	–9.5000	67.6875	–428.6876			
C	3								1.	– 14.2500	135.3750			
	4									1.	– 19.0000			
	5										1.			
	6						1.	–3.7500	14.0625	– 52.7344	197.7539			
	1						.324588	–2.456529	– .580849	–1.513803	.175119	–4.051474	1.	–4.051474
	2							.517164	.244568	.956086	–.147469	1.570349	.125	.196294
D	3								– .025744	– .201281	.046569	– .180456	.015625	– .00282
	4									.014125	–.006536	.007589	.001953	.0000148
	5										.000344	.000344	.000244	.0000000839
	6						.324588	–1.939365	– .362025	– .744873	.068027	–2.653648		

NOTE ON THE COMPARATIVE RELIABILITY OF THE STANFORD-BINET BELOW THE AGE OF SIX

DOROTHY DURLING

Bureau of Mental Health, Pennsylvania Department of Welfare

Comparatively speaking the Stanford-Binet below the age of six, having been standardized upon a relatively small number of cases is regarded as less reliable than the upper part of the scale. This paper offers some objective data upon the reliability of this lower section of the Stanford-Binet, which, while it is by no means conclusive is interesting for purposes of comparison. Our method (which we admit is not an entirely adequate method of showing reliability) would seem to indicate that contrary to the general supposition the Stanford-Binet below the age of six is *not* much lower than that of the upper part of the scale.

Our method of showing reliability is the method of correlation between two trials of the same test, that is, of two Stanford-Binet ratings. Two hundred sixteen Stanford-Binet retests were available in the records of the Pennsylvania Bureau of Mental Health and from these those cases having a mental age of less than six were selected ($N = 76$) and the correlation (Pearson r) between first and second tests was computed.

(1) With the retesting coefficient for the entire 216 cases, which include the entire scale from mental age 3 through 18;

(2) With the retesting coefficient reported by Florence Goodenough for the Kuhlmann-Binet in the age level below 6 (300 cases aged 2, 3 and 4 years).[1] This latter comparison was made because of the alleged superior standardization of the Kuhlmann-Binet (as compared with the Stanford-Binet) for the age levels below 6.

[1] Florence Goodenough, Kuhlmann-Binet tests, University of Minnesota Press, Minneapolis, 1927, p. 62 ff.

	RETESTING COEFFICIENT	NUMBER OF CASES
Stanford-Binet below mental age 6	.79 ± .02	76
Stanford-Binet—entire scale	.83 ± .02	216
Kuhlmann-Binet below mental age 6	.82 ± .01	300

Taking the figures at their face value the retesting coefficient for the Stanford-Binet below the age of 6 compares favorably with the retesting coefficient of the Stanford-Binet as a whole and for the Kuhlmann-Binet in the age level below 6.

It is understood that our method of securing reliability is only an indirect one. Better methods of indicating reliability are:

(1) Coefficient of correlation between two forms of a test.

(2) Correlation between two random halves of the test, taking odd and even items, and corrected by the Spearman-Brown formula.

We say that our method of securing reliability is only an indirect one. Further than that, on account of the interval between the two tests certain factors aside from technical reliability, such as irregularities in the curve of mental growth, may affect the size of the correlation coefficient. Nevertheless technical reliability is such a preponderant factor controlling the retesting coefficient that we have arbitrarily chosen to disregard the additional factors which accrue in the longer intervals between testing.

Coefficients of correlation must always be interpreted in terms of the heterogeneity of the groups studied. Heterogeneity tends to result in a spurious raising of the coefficient of correlation over and above a real relation between the two variables. Therefore, it should be pointed out that our group below mental age 6 is more heterogeneous as to chronological age than the group studied by Goodenough. Below is a table of the ages for our group:

AGE FIRST EXAMINATION	NUMBER OF CASES
3 yrs. 0 mos. to 3 yrs. 11 mos.	2
4 yrs. 0 mos. to 4 yrs. 11 mos.	5
5 yrs. 0 mos. to 5 yrs. 11 mos.	11
6 yrs. 0 mos. to 6 yrs. 11 mos.	18
7 yrs. 0 mos. to 7 yrs. 11 mos.	19
8 yrs. 0 mos. to 8 yrs. 11 mos.	9
9 yrs. 0 mos. to 9 yrs. 11 mos.	4
10 yrs. 0 mos. to 10 yrs. 11 mos.	3
11 yrs. 0 mos. to 11 yrs. 11 mos.	2
12 yrs. 0 mos. to 12 yrs. 12 mos.	1
13 yrs. 0 mos. to 13 yrs. 13 mos.	0
14 yrs. 0 mos. to 14 yrs. 14 mos.	1
15 yrs. 0 mos. to 15 yrs. 15 mos.	1
Over 16	1

Below is a table showing interval between tests:

INTERVAL BETWEEN TESTS	NUMBER OF CASES
1 yr. to 1 yr. 11 mos.	35
2 yrs. to 2 yrs. 11 mos.	23
3 yrs. to 3 yrs. 11 mos.	7
4 yrs. to 4 yrs. 11 mos.	4
5 yrs. to 5 yrs. 11 mos.	6
6 yrs. to 6 yrs. 11 mos.	1

NOTES AND NEWS

Dr. Karl M. Dallenbach, who has been Visiting Professor of Psychology at Columbia University for the past two years, returns to Cornell University this fall as Professor of Psychology.

The Board of Directors of the National Committee on Prisons and Prison Labor, at a recent meeting appointed Dr. Jacob L. Moreno, Consultant in Psychiatry and Criminology.

The American Association for the Advancement of Science held its summer convention in Syracuse, New York, June 20–25. On Wednesday, June 22, Section I, the section on psychology, conducted a symposium on Industrial Psychology, the theme of which was along the line of possible contribution of psychology to industry during a period of business depression. Among the speakers were Dr. Paul S. Achilles, Secretary of the Psychological Corporation, and Mr. W. E. Parker, psychologist for the Employment Center of Rochester, New York.

The Southern California Branch of the American Eugenics Society held its last regular meeting of the season at the Los Angeles Public Library on Monday, June 13. Dr. Frank C. Reid of the Human Betterment Foundation, Pasadena, spoke on the subject, "Quality or Quantity in Race Suicide?" An interesting review of the new book by F. C. S. Schiller on "Social Decay and Eugenical Reform" was presented by Dr. Paul Popenoe. The members of the Society were invited by Dr. Norman Fenton, Director of the State Bureau of Juvenile Research, to visit Pacific Colony on July 12. A special school for research and social workers will be held at Pacific Colony this summer.

The Carnegie Foundation for the Advancement of Teaching, in its current "Annual Review of Legal Education," reports a total expenditure upon the study of legal education and cognate matters during a period of a little over eighteen years of $238,630. Besides the "Annual Review" the Foundation has published several volumes on various phases of legal education, among which are "The Common Law and the Case Method in American University Law Schools" (1914); "Justice and the Poor" (1919, 1921, 1924); "Training for the Public Profession of the Law" (1921); "Present-Day Law Schools in the United States and Canada" (1928). Copies of the publications may be had

without charge upon application to the office of the Foundation, 522 Fifth Avenue, New York City.

United States Commissioner of Education, William John Cooper, recently announced that committees composed of leading citizens in each state have been named to work with the National Survey of School Finance. These committees include legislators, school officials, university professors and other prominent citizens. It was deemed advisable by the Survey's Board of Consultants to invite participation by these "contact" groups who might throw light on the problems of financing the schools which confront each state. Leaders selected in each state will be informed of the progress and advance findings of the nation-wide expenditure study and will be asked for advice from time to time by the survey research staff. Names of committeemen appointed to date for each state may be obtained from the Commissioner of Education, U. S. Department of the Interior, Washington, D. C.

On January 8, 1932, was held the fifth meeting of the Executive Committee of the International Bureau of Education in the Hotel de Ville, Geneva. All the Governments, members of the Bureau, were represented at this meeting. Dr. Piaget is the Director of the Bureau.

The Ohio College Association has recently published several new Bulletins (see New Books list on later page) among which may be mentioned the following: "The Work-Limit Method *versus* the Time-Limit Method in the Construction and Administration of College Ability Tests," by B. E. Workman, Ohio University; "The Validation of the Rating Scales Used with Candidates for Admission to Oberlin College," by L. D. Hartson, Oberlin College. The Committee on Intelligence Tests for Entrance has also recently published its Seventh Annual Report in which is a detailed outline of the work accomplished during the year 1931–32. The membership of this Committee is as follows: Dr. H. A. Toops, Chairman; Dr. L. D. Hartson, Oberlin College; Dr. L. A. Thompson, Ohio Wesleyan University; Dr. W. S. Guiler, Miami University; Dr. M. N. Neuberg, Wittenberg College; Dr. O. F. Mathiasen, Antioch College; and James P. Porter, Ohio University.

The Committee on Diagnosis and Retraining of the Employment Stabilization Research Institute at the University of Minnesota announces that the Carnegie Corporation has appropriated $75,000 to enable the Committee to complete its two-year research program. Psychologists interested in receiving research bulletins should write to the University Press, University of Minnesota, Minneapolis.

The Commonwealth Fund, 41 East 57th Street, New York, recently published the thirteenth Annual Report of its work for the year ending September 30, 1931. Appropriations amounting to $2,232,261.91 are announced for 1931 by the President, Edward S. Harkness, which is the largest sum disbursed by the Fund in any year since it was founded in 1918 by Mrs. Stephen V. Harkness. More than two-thirds of this total was devoted to the promotion of physical and mental health. The number of full-time child guidance clinics in the United States has increased from 20 to 30 in the past two years, and there are now 232 clinics in which children are given care by a team consisting of psychiatrist, psychologist, and psychiatric social worker. Realizing the great need for trained workers, the Fund provided 74 scholarships for psychiatrists and psychiatric social workers in 1931. Besides appropriating $219,593 for the support of the Institute for Child Guidance in New York the Fund also met the costs of a child guidance clinic in London and made a grant to the National Committee for Mental Hygiene toward the cost of a nation-wide survey of psychiatric education in the United States.

The *Journal* is in receipt of the first three issues of a publication by the School Psychotechnical Laboratory of the District Institute for the Investigation of Children in Leningrad. These are accompanied by a description of the work of the Laboratory as a whole. "The Laboratory in its vocational guidance work is not only concerned in ascertaining the vocational capacities of the boys and girls in their teens but it is considered as most important, that it would influence the active formation of their vocational aptitudes in accordance with the interests of the building of socialism, during the whole of their stay at schools of general knowledge," states Professor L. P. Boltenof, Director of the Laboratory.

The Third International Congress of Eugenics will be held at the American Museum of Natural History, 77th Street and Central Park West, New York City, from August 21–23, 1932. An exhibit of materials showing the history, content, and present researches and trends of eugenics will be held in Educational Hall from August 22 to September 22. The Secretary of the Congress and chairman of the Exhibits Committee is Harry H. Laughlin, Cold Spring Harbor, Long Island, N. Y. Those who desire to place material in this exhibit should communicate with him.

BOOK REVIEWS

EMMET, BORIS. *Department Stores, Recent Policies, Costs and Profits.* Stanford University Press, Stanford University, California. 1930. xxiii + 214 pp. $4.00.

In recent years greater attention has been given by accountants and business men generally to the subject of merchandising or distribution costs. Professor Emmet's book is, therefore, very timely. The principles brought out in this book apply to the whole field of distribution, including Chain Stores, Wholesale and Retail Establishments as well as Department Stores. The illustrations are, however, drawn from the Department Store business and the standards developed are useful chiefly in Department Store management.

Following a discussion of Present Conditions in the department store business, covering the subject of retail distribution, general policies, management and personnel, store merger and competition, the author analyzes the trends in department store costs, earnings, returns and allowances, mark-downs, collections and merchandise turnover for the period 1922–1929. In part three of the book he gives certain detailed statistical information by departments concerning operating results, customer returns, mark-downs, stock turnover, sales and stocks.

The author has included in the text numerous charts in addition to the statistical tables, statements and other illustrations mentioned above. The Appendix contains an essay on the meaning of retail accounting and a complete description of the organization and activities of the National Retail Dry Goods Association.

The book is well written and should serve as a challenge to the department store manager of to-day to improve his accounting and management methods so that he may be in position to meet successfully changing conditions in the field of merchandising.

OSCAR S. NELSON,
University of Pennsylvania.

GOW, CHARLES R. *Elements of Human Engineering.* Macmillan, New York, 1932. 169 pp.

This book is a course of lectures given to a class in Humanics at the Massachusetts Institute of Technology. A similar volume was published a few years ago and the present work evidently represents a development of that same course through a period of years. It constitutes a very distinct improvement over the earlier effort. The first was largely inspirational, preachy, and fixed up with unusually tricky titles for the various chapters. The present discussion is more straightforward and there is no offensive, figurative language. Each chapter takes some

single point, elaborates it and then gives rather numerous illustrations and anecdotes to substantiate that point. These illustrations are well-chosen and many of them from the engineering field in which the author has had his principal experience.

A mere mention of the content of these chapters will indicate the general nature of the book. The importance of favorable, initial impressions is stressed. The pigmy and the giant introduces a discussion of the big consequences of small things as in the instance where the pigmy crawls through the keyhole and opens the door for the giant. Obedience makes for a smooth running organization. Persons should not take themselves too seriously, should not let anger dominate their activities and should not be unduly critical all the time. Common sense is distinguished from emotion and it is important to concentrate on the fundamental issues. A few points are made regarding the efficiency of labor as related to cost, buying power, etc. The present change in attitude toward workers is mentioned and such difficulties as stereotyping of output, readjustment of piece-work rate and fear of unemployment. The importance of getting the right man on the right job is brought out, although no techniques for doing so are presented. There is a plea for ethics and sportsmanship. Cheerfulness and optimism will eliminate friction. Success comes usually to those who desire it. Happiness is a by-product of a well adjusted life. The last chapter is essentially an autobiography.

The work is not particularly psychological in character, although there is occasional mention of psychological principles. As far as the reviewer can see there are no psychological fallacies committed. It is a type of common-sense presentation which would probably be valuable for the group for whom it was designed. It includes considerable good advice for a young man leaving college or some similar institution. The author is to be commended for not playing it up as an especially psychological treatise. It is a book which would be satisfactory to put into the hands of any young college graduate. It would certainly do no harm, might give him some ideas that would be of value and might furnish him just a bit of inspiration.

H. E. BURTT,
Ohio State University.

WILLIAMS, FRANKWOOD E., M.D. *Adolescence, Studies in Mental Hygiene.* Farrar and Rinehart, 1930. xi + 279 pp.

The author from his experiences as a psychologist and as a doctor of medicine discusses the topic in an interesting and non-technical style. The general theme is that people mature physically but they do not mature emotionally. The business man, the lawyer, the judge, the professional man appear as adults but they have childish emotional reactions.

In the first lecture he asks, Who is an adult? One can no more make himself mature than he can turn on his pituitary gland or raise himself by his own bootstraps. Three criteria of maturity are offered: one who is able to see objects, persons, acts, in the terms of what they are, cleared of all infantile symbolic investments; one who is under no compulsion either to do or not to do, but who is free to act or not to act according to the realities of the situation; and one who is able to adjust to an unalterable situation with a minimum conflict. Some of the signs by which we may recognize immaturity are: parents who are embarrassed by the sex questions of their children, men who would rather mother their children than to provide for them, husbands who are more devoted to their mothers than to their wives, husbands who are not understood, individuals who are afraid to ask for a raise, women who do not believe that women are inferior to men but who feel so and act as though they were.

In Idealistic But Infected Homes there is a vicious circle. Emotionally immature parents infect with their own infantile and childish reactions children who are thereby prevented from attaining a healthy emotional maturity themselves and another group of parents is produced as unhappy and inadequate as the first. What are parents for? To watch the unfolding of the child and not to mold it. Innocent But Dangerous Parents do not know how to bring about a nice adjustment in the child between parental authority and love and often cause emotional cramps. The child does not develop from a dependent, egocentric, irresponsible state to one of independence, altruism and responsibility. "Parents do more than determine through their germ plasm the color of their children's hair and eyes. They determine how their children will get on with their husbands and wives and their own children." Can youth be coerced? Yes, but why should it be coerced? Because we have no faith in it. Then we set about making the youth into the same failures we have been. Youth is striving for just two things: love in the widest sense and a feeling of security. In Adolescent Confronting the World: His Two Problems are set forth as emancipation from home and adequate heterosexual adjustment. This is one of the best chapters in the book and is likely to cause disgust in the possessive father and mother who are not willing for their children to become independent. The adolescent boy is keen for advice, but he resents, and rightfully so, any appeal for good conduct on the ground of love for his mother. Emancipation from home means the psychological freeing of one's self from childish bonds.

In the last section of the book the author draws a distinction between mental hygiene as an organized social movement and mental hygiene as an art in the application of knowledge derived from certain basic sciences to the maintenance of individual mental health. Both are necessary in the community. Some of the responsibility of the commu-

nity is indicated in the facts that in the last twenty-four hours 7,000 children were born in this country, and last year 70,000 other children dropped out of the race of life and were admitted to the hospitals for the insane, and there are 350,000 people in the United States within from five days to five years of the hospitals.

Thoughtful readers will be benefited and will enjoy this book.

J. R. GENTRY,
Ohio University.

CLYDE HISSONG. Educational Psychology Monographs, No. 30. *The Activity Movement.* Warwick and York, Inc., 1932.

That the principles of Progressive Education were centuries in evolving, Clyde Hissong, in his "The Activity Movement," attributes to the fact that educationists from Rousseau to Dewey were handicapped by the lack of fundamental facts and principles of human nature upon which to base their philosophies. Following a careful analysis of the theories of Rousseau, Herbart, G. Stanley Hall, Thorndike, and modern biology, he arrives at the conclusion that the last named "points toward a unitary organism interacting within the environment" as being in harmony with the principles of the activity movement.

Thus far all is clear. But his treatment of the practices of progressive schools leaves one wondering just what his position *is* in regard to the activity movement. While recognizing numerous worthy contributions that are in harmony with the unitary concept he also points out certain omissions that endanger the success of the movement. These omissions as he sees them are: (1) the absence of a pervading program; (2) the absence of an organized knowledge background; (3) neglecting the formation of flexible habits; and (4) no directing of the child to see the interrelation between situations. Besides giving attention to the meeting of these requisites he further suggests that the teacher give emphasis to vocational facility, to intellectual interests and to the development of the social outlook—the relationship of the individual to life as a whole.

In the initial chapter the author pointed out the distinction between the radical activity schools as represented by The School of Organic Education and the more conservative activity schools as represented by the Lincoln School. But in his final analysis of the contributions and omissions of the activity movement he seems inclined to believe that practically all activity schools are as radical as Marietta Johnson's. One wonders if the author has been intimately associated with children and teachers working and living in an activity school or if he has only just read about them!

EPSIE YOUNG,
Ohio University.

NEW BOOKS AND PAMPHLETS RECEIVED

Books and pamphlets for review should be sent to James P. Porter, Editor of THE JOURNAL OF APPLIED PSYCHOLOGY, Ohio University, Athens, Ohio.

The Behavior of Young Children. III. ETHEL B. WARING AND MARGUERITE WILKER. Charles Scribner's Sons, New York City, 1932. 198 pp.

Elements of Human Engineering. CHARLES R. GOW. Macmillan Company, New York City, 1932. $1.60. 169 pp.

An Evaluation of the Minnesota Rating Scale for Home Economics Teachers. CLARA M. BROWN. University of Minnesota Press, Minneapolis, Minn. $.50. 29 pp.

An Experimental Study of Pitch Recognition. LAURENCE A. PETRAN. Psychological Monographs, Vol. XLII, No. 6. Psychological Review Company, Princeton, N. J., 1932. 124 pp.

Influence of Stuttering on the Personality. WENDELL JOHNSON. University of Iowa, Studies in Child Welfare, Vol. 5, No. 5, 1932. University of Iowa, Iowa City, Iowa. 140 pp.

The Medical Value of Psychoanalysis. FRANZ ALEXANDER. W. W. Norton & Company, New York City, 1932. $2.75. 247 pp.

La Methode des Tests. RENE NIHARD. Les Editions du Cerf. Juvisy, Seine-et-Oise, France. 235 pp.

Ohio College Association Bulletins, Ohio State University, Columbus, Ohio:

The Contribution Made by "Difficulty of Course" to the Index for Predicting Freshman Scholarship. L. D. HARTSON, Oberlin College. Bulletin No. 71.

Empirical Psychology and the "Generalized Regression" Equation. HERBERT A. TOOPS, Ohio State University. Bulletin No. 81.

The Work-Limit Method Versus the Time-Limit Method in the Construction and Administration of College Ability Tests. B. E. WORKMAN, Ohio University. Bulletin No. 79.

Predicting Soldiers' School Marks: A Problem in Selection of Tests. HERBERT A. TOOPS AND ELMER B. ROYER, Ohio State University. Bulletin No. 80.

1931–32 (Seventh) Annual Report of the Committee on Intelligence Tests for Entrance. Bulletin No. 82.

On the Validity of Two Multiplied Traits as Compared with Two Added Traits, Fundamental Equality of the Units of Measurement Within

the Given Scale of the Several Variables. HERBERT A. TOOPS. Bulletin No. 85.

The Validation of the Rating Scales Used with Candidates for Admission to Oberlin College. L. D. HARTSON, Oberlin College. Bulletin No. 83.

Partial List of Research Projects at Present Under Way or in Prospect in American Institutions of Higher Education in the Fields of (1) Student Personnel, (2) Collegiate Instruction and Curricula, (3) Improvement of Methods of Testing, (4) Administration. United States Department of the Interior, Office of Education, Washington, D. C. 54 pp.

Proceedings of the First International Congress on Mental Hygiene. Vols. I and II. Edited by FRANKWOOD E. WILLIAMS. International Committee for Mental Hygiene, New York City, 1932.

Psicologia Juridica. EMILIO MIRA Y LOPEZ. Salvat Editores, S. A. Barcelona, Spain, 1932. 270 pp.

A Scientific Approach to Labor Problems. ADELBERT FORD. McGraw-Hill Book Company, New York City. $4.00. 446 pp.

Secret Ways of the Mind. W. M. KRANEFELDT. Translation by Ralph M. Eaton. Henry Holt & Company, New York, 1932. 188 pp.

A Series of Experiments on the Learning of French Vocabulary. LOUISE C. SEIBERT. Johns Hopkins University Studies in Education, No. 18. Johns Hopkins Press, Baltimore, Md. 106 pp.

Student Self-Support at the University of Minnesota. JAMES B. UMSTATTD. University of Minnesota Press, Minneapolis, Minn., 1932. $2.50. 205 pp.

Twenty-fifth Anniversary of the Vineland Laboratory. EDGAR A. DOLL, Editor. The Training School at Vineland, N. J., Department of Research, 1932. 135 pp.

Workers' Emotions in Shop and Home. REXFORD B. HERSEY. University of Pennsylvania Press, 3438 Walnut Street, Philadelphia, Pa., 1932. $3.00. 441 pp.

Directory of American Psychological Periodicals

(Continued)

Psychological Abstracts—Eno Hall, Princeton, N. J. American Psychological Association. Subscription $6.00. 700 pages ann. Edited by W. S. Hunter. Monthly. Abstracts of psychological literature. Founded 1927.

Journal of General Psychology—Worcester, Mass.: Clark University Press. Subscription $14.00 per annum; $7.00 per volume. 1000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

Journal of Social Psychology—Worcester, Mass.: Clark University Press. Subscription $7.00. 500 pages annually. Edited by John Dewey, Carl Murchison, and an international cooperating board. Quarterly. Political, racial, and differential psychology. Founded 1929.

Personnel Journal—Baltimore: The Williams & Wilkins Co. Subscription $5.00. 500 pages, approximately. Edited by Walter V. Bingham. Bi-Monthly.

Mental Measurement Monographs—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Edited by Knight Dunlap and Buford Johnson. Published without fixed dates, each number a single research. Founded 1925.

Studies in Psychology and Psychiatry—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Edited by Edward A. Pace. Published without fixed dates, each number a single research. Founded 1926.

Journal of Philosophy—New York: 515 West 116th Street. Subscription $4.00. 728 pages per volume. Founded 1904. Bi-weekly. Edited by F. J. E. Woodbridge, Wendell T. Bush, and H. W. Schneider.

Child Development—Baltimore, Md.: Williams & Wilkins Company. Subscription. $5.00. 400 pages annually. Edited by Buford Johnson. Quarterly. Founded 1930.

Directory of American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.: Cornell University. Subscription $6.50. 624 pages annually. Edited by M. F. Washburn, K. M. Dallenbach, M. Bentley, E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

Pedagogical Seminary and Journal of Genetic Psychology—Worcester, Mass.: Clark University Press. Subscription $14.00 per annum; $7.00 per volume. 1000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Child behavior, animal behavior and comparative psychology. Founded 1891.

Psychological Review—Princeton, N. J.: Psychological Review Company. Subscription $5.50. 540 pages annually. Bi-monthly. General Psychology. Edited by Howard C. Warren. Founded 1894.

Psychological Monographs—Princeton, N. J.: Psychological Review Company. Subscription $6.00 per vol. 500 pp. Ed. by Raymond Dodge. Published without fixed dates, each issue one or more researches. Founded 1895.

Psychological Index—Princeton, N. J.; Psychological Review Company. Subscription $4.00. 300–400 pp. Edited by W. S. Hunter. An annual bibliography of psychological literature. Founded 1895.

Psychological Bulletin—Princeton, N. J.: Psychological Review Company. Subscription $6.00. 720 pages annually. Psychological literature. Monthly. Edited by Edward S. Robinson. Founded 1904.

Archives of Psychology—New York: Columbia University. Subscription $6. 500 pp. ann. Ed. by R. S. Woodworth. Published without fixed dates, each number a single experimental study. Founded 1906.

Journal of Abnormal and Social Psychology—Eno Hall, Princeton, N. J. American Psychological Association. Subscription $5. Foreign, $5.25. Editor, Henry T. Moore. Quarterly. 448 pages annually. Abnormal and Social. Founded 1906.

Psychological Clinic—Philadelphia: Psychological Clinic Press. Subscription $3.00. 288 pages. Ed. by Lightner Witmer. Without fixed dates (9 Numbers). Orthogenics, psychology, hygiene. Founded 1907.

Psychoanalytic Review—Washington, D. C.: 3617 10th St., N. W. Subscription $6. 500 pages annually. Psychoanalysis. Quarterly. Ed. by W. A. White and S. E. Jelliffe. Founded 1913.

Journal of Experimental Psychology—Princeton, N. J. Psychological Review Company. 500 pages annually. Experimental. Subscription $6.00. Bi-monthly. Edited by S. W. Fernberger. Founded 1916.

Journal of Applied Psychology—Lancaster: The Science Press Printing Co. Subscription $5.00. 400 pages annually. Bi-monthly. Edited by James P. Porter, Ohio University, Athens, Ohio. Founded 1917.

Journal of Comparative Psychology—Baltimore: The Williams & Wilkins Co. Subscription $5.00 per volume of 450 pages. Three volumes every two years. Edited by Knight Dunlap and Robert M. Yerkes. Founded 1921.

Comparative Psychology Monographs—Baltimore: Johns Hopkins University Press. Subscription $5. 400 pages per volume. Knight Dunlap, Managing Editor. Published without fixed dates, each number a single research.

Genetic Psychology Monographs—Worcester, Mass.: Clark University Press. Subscription $14.00 per annum; $7.00 per volume. 1000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Monthly. Each number one complete research. Child behavior, animal behavior, and comparative psychology. Founded 1925.

(Continued on Third Cover)